IOWA STATE UNIVERSITY
Digital Repository

Retrospective Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

2006

Chemical reagents in the study of biomolecules

Laura-Mirela Dutcă
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

 Part of the Biochemistry Commons

Recommended Citation
Dutcă, Laura-Mirela, "Chemical reagents in the study of biomolecules " (2006). *Retrospective Theses and Dissertations*. 1905.
https://lib.dr.iastate.edu/rtd/1905

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

www.manaraa.com

Chemical reagents in the study of biomolecules

by

Laura-Mirela Dutcă

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemistry (Biomolecular Sciences)

Program of Study Committee
Gloria Culver, Co-major Professor
Jacob Petrich, Co-major Professor
Amy Andreotti
Aaron Sadow
Emily Smith
Keith Woo

Iowa State University

Ames, Iowa

2006

Copyright © Laura-Mirela Dutcă, 2006. All rights reserved

www.manaraa.com

UMI Number: 3244376

UMI

UMI Microform 3244376

Copyright 2007 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

www.manaraa.com

TABLE OF CONTENTS

www.manaraa.com

www.manaraa.com

Chapter I. General introduction

Introduction

Chemists are inspired by the amazing systems present in nature, and biologists are using tmethods developed by chemists to study these systems. The tools of chemistry and biology are combined all the time to probe biological problems. Both chemistry and biology continue to grow and draw inspiration from each other as answers create more questions.

There are many applications of chemical reagents in biology. It is not the purpose of this introduction to review all these applications; some of them will be mentioned to emphasize the relevance of chemical reagents for the study of biological systems. Chemical reagents are an essential complement to all the techniques that biologists have developed to facilitate the selective study of biomolecules in native contexts. Chemical reagents are employed because they are more readily available than many biomolecules (for example enzymes), their properties can be adjusted, they can usually be synthesized in large quantities, their size can be varied, and sometime of course they have useful properties that are not exhibited by natural compounds. One of the main characteristics of synthetic chemistry is control over molecular structure and function. The combined efforts of many chemists lead to the development of sophisticated strategies that make possible the synthesis of a vast number of compounds and the tailoring of their properties, which are huge advantages when designing reagents for certain tasks. It is not an easy task to modulate the properties of proteins, DNA, RNA, or other large biomolecules, and probably this is one of the biggest advantages of chemical reagents.

www.manaraa.com

Many applications of chemical reagents in biology are in the study of structure of biomolecules, and consequently how biomolecules interact with each other. Structure and function are intimately connected, and ultimately, the structure of biomolecules reveals how they perform their function. Chemical reagents are used to determine the primary sequence of biomolecules, analyze their three dimensional structure, and to dissect their interaction with other biomolecules.

Modification of biomolecules with chemical reagents is the main approach mentioned in this thesis and it is widely used in the study of proteins and nucleic acids.[1-3] Site specific chemical modification of proteins is a very useful tool in proteomics. For example mass spectrometric analysis is used in combination with different chemical modifications, or incorporation of isotope-coded affinity tags.[4,5] Chemical labeling with fluorescent probes is employed in analyzing conformational changes in biomolecules.[6-8] Conformational changes in nucleic acids during folding or interaction with other molecules are explored with base-specific chemical probes.[1,9,10] More details about base-specific probing of RNA will be presented in Chapter V.

Cleavage of biomolecules is a useful chemical modification, and since this approach is used prominently with proteins and RNA in this thesis, some examples of applications of cleavage reagents are discussed. Cleavage of biomolecules is an important step in the determination of their primary sequences.[11] Footprinting and folding studies use changes in the cleavage pattern to obtain structural information.[12-15] For proteins, cleavage reagents are also used in other applications. Protein digests are used in proteomics for peptide mapping.[16] Fusion tags are removed by site-specific cleavage from engineered proteins to convert them to their native form.[17] Fragments of natural

www.manaraa.com

proteins are ligated with other natural or synthesized fragments in protein semisynthesis.[18]

Enzymes and ribozymes are highly selective and efficient catalysts, but chemical reagents have a prominent place in the study of biological systems. Indeed, many of the reactions performed by chemical reagents are performed better by enzymes and ribozymes, but size, reaction conditions or even availability can make the utilization of chemical reagents more suitable.

Dissertation organization

The first chapter of this dissertation presents some applications of chemical reagents in biological sciences. This dissertation is composed of two different parts. The first part presents the work that I have done under the supervision of Dr. Nenad Kostić, on *cis*-$[Pt(en)(H_2O)_2]^{2+}$ as a peptide cleavage reagent. Chapter II is a review on Pt(II) and Pd(II) complexes as synthetic proteases. Our work on *cis*-$[Pt(en)(H_2O)_2]^{2+}$ produced three papers,[19-21] one of which is included as Chapter III, and was published in Inorganic Chemistry.[21] The other two,[19,20] which resulted from my collaboration with Nebojsa Milović, will not be reproduced here since they are include in his thesis. In one of the papers published with Nebojsa Milović, the activity of *cis*-$[Pt(en)(H_2O)_2]^{2+}$ as a peptide cleavage reagent is described, along with its selectivity and mechanism of cleavage, and it was published in *Chemistry-A European Journal.*[19] Our results for the combined use of platinum(II) and palladium(II) complexes as selective cleavage reagents were published in *Inorganic Chemistry.*[20] The third paper, which is included here, is a collaboration with Dr. Nicola Pohl's group on the influence of different types of irradiation on cleavage of peptides and proteins by *cis*-$[Pt(en)(H_2O)_2]^{2+}$. Kwang-Seuk Ko helped in the experiments

www.manaraa.com

involving cleavage under microwave irradiation. Dr. Nicola Pohl suggested the cleavage of peptides and proteins under microwave irradiation. Dr. Nenad Kostić got the suggestion to try cleavage under ultraviolet irradiation at a conference. Chapter IV summarizes our findings on Pt(II) reagents as synthetic proteases.

My work with Dr. Gloria Culver focused on protein-RNA interactions during assembly of the small subunit of the prokaryotic ribosome. Chapter V is a general introduction on the small 30S subunit of the ribosome and methods to explore RNA-protein interactions, including techniques that involve chemical reagents. Chapter VI presents our results on the study of interaction of 16S rRNA with primary binding ribosomal proteins (r-protein) at different temperatures. The interaction of individual primary r-proteins with 16S rRNA is studied by base specific chemical footprinting at low and high temperatures. For this study, the experiments involving r-protein S15 were performed by Indu Jagannathan, and some of the experiments involving r-protein S8 were done by Joel Grondek. We also used directed hydroxyl radical probing to study changes in the architecture of 16S rRNA surrounding S20 during assembly of the 30S subunit. A new approach in the exploration of RNPs of different complexities by directed hydroxyl radical probing and our preliminary results are presented in chapter VII. A summary of our results obtained in the study of RNA-protein interactions and some future directions constitute the last chapter.

References

1. Stern, S.; Moazed, D.; Noller, H. F., *Methods Enzymol.* **1988,** 164, 481-9.

2. Christiansen, J.; Garrett, R., *Methods Enzymol.* **1988,** 164, 456-468.

www.manaraa.com

3. Lundblad, R., *Chemical Reagents for Protein Modification*. 3rd ed.; CRC Press: Boca Raton, 2005.

4. Leitner, A.; Lindner, W., *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2004,** 813, (1-2), 1-26.

5. Tao, W. A.; Aebersold, R., *Curr. Opin. Biotechnol.* **2003,** 14, (1), 110-8.

6. Walter, N. G., *Methods* **2001,** 25, (1), 19-30.

7. Klostermeier, D.; Millar, D. P., *Methods* **2001,** 23, (3), 240-54.

8. Vigano, C.; Manciu, L.; Ruysschaert, J. M., *Acc. Chem. Res.* **2005,** 38, (2), 117-26.

9. Moazed, D.; Stern, S.; Noller, H. F., *J. Mol. Biol.* **1986,** 187, (3), 399-416.

10. Moore, P. B., The RNA folding problem. In *The RNA World*, Gesteland, R. F.; Cech, T. R.; Atkins, J. F., Eds. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1999; pp 381-401.

11. Croft, L. R., *Handbook of Protein Sequence Analysis*. 2nd ed.; Wiley: Chichester, U. K., 1980.

12. Heyduk, T.; Baichoo, N.; Heyduk, E., *Met. Ions Biol. Syst.* **2001,** 38, (Probing of Proteins by Metal Ions and Their Low-Molecular-Weight Complexes), 255.

13. Hubbard, S.; Beynon, R. J., *Proteolysis of Native Proteins as a Structural Probe*. Oxford University Press: New York, 2001; p 233.

14. Joseph, S.; Noller, H. F., *Methods Enzymol.* **2000,** 318, 175-90.

15. Culver, G. M.; Noller, H. F., *Methods Enzymol.* **2000,** 318, 461-75.

16. Thomas, J. J.; Bakhtiar, R.; Siuzdak, G., *Acc. Chem. Res.* **2000,** 33, (3), 179-187.

www.manaraa.com

17. Thorner, J.; Emr, S. D.; Abelson, J. N.; Editors, *Applications of Chimeric Genes and Hybrid Proteins, Part A: Gene Expression and Protein Purification. [In: Methods Enzymol., 2000; 326]*. 2000; p 617.

18. Wallace, C. J. A., *Protein Engineering by Semisynthesis*. CRC Press: Boca Raton, FL, 2000.

19. Milovic, N. M.; Dutca, L. M.; Kostic, N. M., *Chemistry* **2003,** 9, (20), 5097-106.

20. Milovic, N. M.; Dutca, L. M.; Kostic, N. M., *Inorg. Chem.* **2003,** 42, (13), 4036-45.

21. Dutca, L. M.; Ko, K. S.; Pohl, N. L.; Kostic, N. M., *Inorg. Chem.* **2005,** 44, (14), 5141-6.

www.manaraa.com

Chapter II. Chemical reagents as selective protein cleavers - Pt(II) complexes

Introduction

Selective proteolysis can be achieved with both enzymes and synthetic reagents.[1] While enzymes have advantages, as being fast, selective and effective in mild conditions, sometime they contaminate the biological samples, and may lack the required selectivity or functionality in the desired conditions. As an alternative for enzymatic digestion, chemical reagents are an important tool for selective cleavage of proteins. However, developing new protein cleavage reagents is not a trivial task. The peptide bond is extremely unreactive toward hydrolysis and even nonselective cleavage is hard to achieve. Under standard conditions, (room temperature, pH 4-8) the half-life for hydrolysis of a simple peptide is 500-1000 years.[2] A few chemical reagents are available for cleavage of proteins,[1] but new chemical reagents with improved efficiency and adjustable selectivity are highly desired. A wide range of proteolyic reagents are necessary. Some applications require long fragments, while other require short ones. In some cases, it is necessary to obtain an unmodified fragment, while sometimes it is desirable to have modified termini. The conditions in which the cleavage takes place also differ; sometimes neutral pH is required, while other applications necessitate acidic pH or the presence of detergents.

Transitional metal complexes are emerging as selective proteolytic reagents.[3] Coordination compounds of Pd(II) and Pt(II) promote hydrolytic cleavage of amide bonds in peptides and proteins.[4-18] Because, Pd(II) ions are more labile than Pt(II) ions,

www.manaraa.com

many studies focused on Pd(II) complexes, to take advantage of their higher reactivity. Recently, we have shown that Pt(II) complexes (with ethylenediamine or 1,3-bis(methylthio)propane as ligands), can be as efficient as Pd(II) complexes in the cleavage of peptide bonds, albeit with different selectivity.[14,15,18]

In this chapter, a short summary of our findings on the cleavage of peptides and proteins by complexes of Pd(II) and *cis*-$[Pt(en)(H_2O)_2]^{2+}$ is presented.

Pt(II) and Pd(II) complexes. Requirements for efficient peptide bond cleavage

Ideally, reagents that cleave peptide bonds should be selective, reactive and removable. In the case of Pt(II) and Pd(II) complexes, the selectivity of reagents is given by anchoring to the metal ion of only some of the amino acid side chains. After selective binding, the metal ion interacts with the target peptide bond and facilitates cleavage. In order to be hydrolytically active, complexes need at least two weakly coordinated ligands, such as water, that can be easily substituted. For charge balance, noncoordinating counterions such as perchlorate ion or tetrafluoroborate are necessary. Coordination of the amide nitrogen to the metal ion strengthens the peptide bond, while coordination of the amide oxygen weakens it.[19] When a metal ion binds the amide oxygen, it enhances the partial positive charge on carbon and stabilizes the tetrahedral intermediate for hydrolysis.[19] Thus, the metal ion should not displace the proton from the NH of the amide bond, but it should interact with the oxygen. The other possible scenario is to deliver a water molecule to hydrolyze the amide bond, but based on the studies to date, the first scenario is the most plausible.[16]

www.manaraa.com

Pd(II) complexes as selective peptide bond cleavers

Many complexes of Pd(II) were studied,[6,8,10-12,20-24] and the most efficient cleavage reagent proved to be $[Pd(H_2O)_4]^{2+}$.[12] In weakly acidic aqueous solutions, cleavage by Pd(II) complexes takes place at the second amide bond upstream from histidine and methionine residues, that is, the X-Y bond in the sequence segments X-Y-His-Z and X-Y-Met-Z, in which X, Y, and Z are any noncoordinating residues (Chart 1). When the pH is raised from mildly-acidic to neutral, cleavage takes place only in X-Pro-His-Z and X-Pro-Met-Z sequences, where the Y residue is praline.[16] In a different approach that does not involve binding of an aminoacid side chain to the metal ion, but utilizes a host guest interaction between cyclodextrin and an aromatic aminoacid, it was shown that a Pd(II)-cyclodextrin conjugate can cleave selectively the X-Pro bond in an X-Pro-Phe sequence.[25]

Pd(II) complexes cleave at methionine, and hystidine,[26] and they bind the N-terminus of peptides or proteins, but this binding is not conducive to cleavage.[13] The initial anchoring of Pd(II) ion facilitates the interaction with the upstream peptide bond, toward the N-terminus of the peptide (Figure 1, complex 1). The anchored Pd(II) ion facilitates deprotonation of the amide NH group of the peptide bond upstream from the anchor, and binds the nitrogen atom of the resulting amidate anion, subsequently inhibiting hydrolysis of this peptide bond (Figure 1, complex 2). Pd(II) is considered the most efficient metal ion in displacing the proton of the amide nitrogen, with a pK_a of ~ 2 for the first bond and ~4 for the second peptide bond in triglycine.[26-28] Since cleavage takes place at pH~2, the Pd(II) ion is not capable to displace the proton from the NH of the second amide bond (pK_a ~4) upstream from the anchor, and promotes its hydrolysis,

www.manaraa.com

probably by interaction with the amide oxygen.[11,16] If other Pd(II) complexes as *cis*-$[Pd(en)(H_2O)_2]^{2+}$ (en is ethylenediamine) are used for cleavage of peptides and proteins, a lag time is observed. This delay arises from the relatively slow displacement of the bidentate ethylenediamine ligand by the donor atoms in the peptide.[12]

cis-$[Pt(en)(H_2O)_2]^{2+}$ as selective peptide bond cleaver

Pt(II) complexes, and in particular *cis*-$[Pt(en)(H_2O)_2]^{2+}$ became the focus of our study as protein cleavage reagents by accident. Due to the high affinity of Pt(II) for sulfur ligands, Pt(II) complexes were considered suitable reagents for blocking cleavage by Pd(II) complexes at methionine. However, instead of preventing cleavage, they facilitated hydrolytic cleavage of peptide bonds with a different selectivity than Pd(II) complexes. In collaboration with Nebojsa Milović, the selectivity and the mechanism of cleavage by Pt(II) complexes were explored.[15] A short summary of our findings on the cleavage of peptides and proteins by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ is presented. The complex of Pt(II) with 1,3-bis(methylthio)propane instead of en as a ligand, was also used as a cleavage reagent.[18] This complex has the same selectivity as *cis*-$[Pt(en)(H_2O)_2]^{2+}$, but cleaves with a slower rate.[14, 18] In my work *cis*-$[Pt(en)(H_2O)_2]^{2+}$ was used, and subsequently this complex will be the focus of this section.

The complex *cis*-$[Pt(en)(H_2O)_2]^{2+}$ cleaves exclusively the Met-Z bond in peptides and proteins in weakly acidic solutions[14](Chart 1). Proximity to the peptide bond is achieved by anchoring to a side chain, and consequently the selectivity of cleavage is governed by the selectivity of the binding to methionine side chains (Figure 2). In peptides and proteins *cis*-$[Pt(en)(H_2O)_2]^{2+}$ can also bind the side chain of cysteine, beside the aforementioned methionine side-chain. Until now, no studies regarding cleavage

www.manaraa.com

directed by cysteine were reported and all the peptides and proteins in which cleavage by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ was performed lacked cysteines. After binding of the Pt(II) reagent to the side chain of methionine and replacement of one of the water ligands by the sulfur atom, the Pt(II) ion will interact with the peptide bond downstream from the side chain.[14] The active complex is formed (Figure 2, complex 1), and hydrolytic cleavage of the first peptide bond downstream from the anchor will take place by interaction with the amide oxygen.[14] The ethylendiamine ligand will stay bound to the Pt(II) ion all the time, in contrast to the case of Pd(II)[12], as it was shown by NMR.[12,14] After the cleavage takes place, the Pt(II) reagent can be removed and peptide fragments with unmodified termini are obtained. The cleavage takes place at a pH lower than 2.5.[14] If the pH is higher than 2.5, the Pt(II) ion will displace the proton of the amide NH group from the peptide bond upstream and form an inactive complex (Figure 2, complex 2). When the deprotonation of the amide NH group upstream takes place a stable six-membered ring is formed, while deprotonation of the downstream NH group would form a less stable seven-membered ring. The formation of the hydrolytically active and of the inactive complexes was shown by NMR.[14]

Peptides and proteins were cleaved by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ at different temperatures (40 and 60^{o}C), at a pH of 2.5 in the presence and absence of sodium dodecylsulfate, a common reagent used for solubilization of proteins.[14] Pt(II) and Pd(II) reagents were used in combination to cleave peptides and proteins, showing their applicability. *cis*-$[Pt(en)(H_2O)_2]^{2+}$ was used to obtain long fragments, and $[Pd(H_2O)_4]^{2+}$ was used for peptide mapping by mass spectrometry.[15]

www.manaraa.com

Different selectivities of Pd(II) and Pt(II) reagents as peptide bond cleavers

Pd(II) and Pt(II) compounds have been studied intensively, since they have many applications in chemistry, medicine or as industrial catalysts. Pt(II) and Pd(II) have a d^8 configuration, are diamagnetic and in the majority of their complexes have a square planar geometry.[29] Consequently, they have quite similar chemical properties. There are a few exceptions, some of which are important for their activity as artificial peptidases, since they explain the different proteolytic activity of the two ions. In studies of ligand substitution it was shown that Pt(II) is inert to ligand substitution, while Pd(II) is much more labile. The rates for ligand substitution at Pd(II) are usually 10^5 times higher than those for similar Pt(II) complexes.[29] Both metal ions are classified as "soft" Lewis acids, and they prefer "soft" ligands, such as sulfur donors to hard ligands, such as oxygen donors.[4, 8, 28] Due to its larger size Pt(II) is "softer" than Pd(II).[28]

A key difference between Pt(II) and Pd(II) ions as peptide cleavage reagents is their selectivity in binding amino acid side chains. Pt(II) will not bind the histidine side chain[27] and subsequently will not cleave the adjacent peptide bond, because it prefers softer ligands like the sulfur atom. The second difference is that Pt(II) will cleave a peptide bond without "losing" its ethylenediamine ligand,[14] while Pd(II) has to lose its ligand before effecting cleavage.[12, 13] The pK_a of the amide nitrogen in the presence of Pd(II) or Pt(II) is another important factor in their selectivity as peptide bond cleavers, as it was discussed above.

www.manaraa.com

cis-[Pt(en)(H$_2$O)$_2$]$^{2+}$ versus cyanogen bromide as selective peptide bond cleavers

Cyanogen bromide, one of the widely utilized methionine selective protein cleavers,[30, 31] and c*is*-[Pt(en)(H$_2$O)$_2$]$^{2+}$ have the same selectivity, they both cleave the Met-Z bond.[1,14,15] Even though they are similar in their proteolytic activity and reaction times, there are quite a few differences between them. Cyanogen bromide has to be present in high excess, up to 100 fold and it requires 70% trifluoroacetic acid or 100% acetic acid as solvent[1]. In contrast, c*is*-[Pt(en)(H$_2$O)$_2$]$^{2+}$ is effective at one to one ratio for each methionine residue, and it cleaves at pH lower than 2.5 and temperatures between 40 and 60^{o}C.[14,15] Moreover, cleavage by cyanogen bromide converts methionine to serine lactone, does not cleave Met-Pro bonds, and has quite a few side reactions. Pt(II) reagent does not modify methionine, cleaves at the C-terminus of methionine even when it is followed by proline, and no side reactions are reported for it.

Our studies showed that Pt(II) reagents cleave hydrolytically peptides and proteins. Their properties make them suitable for biochemical applications, such as peptide mapping or generation of long peptide fragments that can be used for chemical ligation or other purposes.

References

1. Walker, J. M., *The Protein Protocols Handbook*. Humana Press: Totowa, NJ, 2002.

2. Radzicka, A.; Wolfenden, R., *J. Am. Chem. Soc.* **1996,** 118, (26), 6105-6109.

www.manaraa.com

3. Sigel, A.; Sigel, H.; Editors, *Probing of Proteins by Metal Ions and Their Low-Molecular-Weight Complexes. [In: Met. Ions Biol. Syst., 2001; 38], Chapters 2-9 and references therein*. Marcel Dekker: New York, 2001.

4. Burgeson, I. E.; Kostic, N. M., *Inorg. Chem.* **1991,** 30, (23), 4299-4305.

5. Zhu, L.; Kostic, N. M., *Inorg. Chem.* **1992,** 31, (19), 3994-4001.

6. Zhu, L.; Kostic, N. M., *Inorg. Chim. Acta* **1994,** 217, (1-2), 21.

7. Parac, T. N.; Kostic, N. M., *J. Am. Chem. Soc.* **1996,** 118, (25), 5946-5951.

8. Parac, T. N.; Kostic, N. M., *J. Am. Chem. Soc.* **1996,** 118, (1), 51-58.

9. Karet, G. B.; Kostic, N. M., *Inorg. Chem.* **1998,** 37, (5), 1021-1027.

10. Parac, T. N.; Kostic, N. M., *Inorg. Chem.* **1998,** 37, (9), 2141-2144.

11. Milovic, N. M.; Kostic, N. M., *Met. Ions Biol. Syst.* **2001,** 38, (Probing of Proteins by Metal Ions and Their Low-Molecular-Weight Complexes), 145-186.

12. Milovic, N. M.; Kostic, N. M., *J. Am. Chem. Soc.* **2002,** 124, 4759-4769.

13. Milovic, N. M.; Kostic, N. M., *Inorg. Chem.* **2002,** 41, (26), 7053-7063.

14. Milovic, N. M.; Dutca, L. M.; Kostic, N. M., *Chemistry* **2003,** 9, (20), 5097-106.

15. Milovic, N. M.; Dutca, L. M.; Kostic, N. M., *Inorg. Chem.* **2003,** 42, (13), 4036-45.

16. Milovic, N. M.; Kostic, N. M., *J. Am. Chem. Soc.* **2003,** 125, (3), 781-788.

17. Dutca, L. M.; Ko, K. S.; Pohl, N. L.; Kostic, N. M., *Inorg. Chem.* **2005,** 44, (14), 5141-6.

18. Stoffregen, S. A.; Griffin, A. K.; Kostic, N. M., *Inorg. Chem.* **2005,** 44, (24), 8899-907.

19. Martin, R. B., *Met. Ions Biol. Syst.* **2001,** 38, 1-23.

www.manaraa.com

20. Chen, X.; Zhu, L.; You, X.; Kostic, N. M., *J. Biol. Inorg. Chem.* **1998,** 3, (1), 1.

21. Milinkovic, S. U.; Parac, T. N.; Djuran, M. I.; Kostic, N. M., *J. Chem. Soc., Dalton Trans.* **1997**, (16), 2771.

22. Korneeva, E. N.; Ovchinnikov, M. V.; Kostic, N. M., *Inorg. Chim. Acta* **1996,** 243, (1-2), 9.

23. Zhu, L.; Qin, L.; Parac, T. N.; Kostic, N. M., *J. Am. Chem. Soc.* **1994,** 116, (12), 5218-5224.

24. Zhu, L.; Kostic, N. M., *J. Am. Chem. Soc.* **1993,** 115, (11), 4566-4570.

25. Milovic, N. M.; Badjic, J. D.; Kostic, N. M., *J. Am. Chem. Soc.***2004,** 126, (3), 696-7.

26. Pettit, L. D.; Bezer, M., *Coord. Chem. Rev.* **1985,** 61, 97-114.

27. Sigel, H.; Martin, R. B., *Chem. Rev.* **1982,** 82, (4), 385-426.

28. Appleton, T. G., *Coord. Chem. Rev.* **1997,** 166, 313-359.

29. Cotton, F. A.; Wilkinson, G., *Advanced Inorganic Chemistry, 5th Edition.* Wiley: New York, 1988.

30. Spande, T. F.; Witkop, B.; Degani, Y.; Patchornik, A., *Adv. Protein Chem.* **1970,** 24, 97-260.

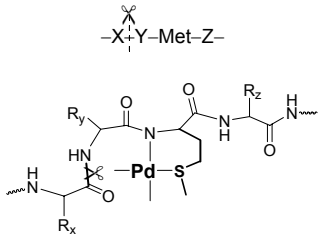
31. Gross, E., *Methods Enzymol.* **1967,** 11, 238-245.

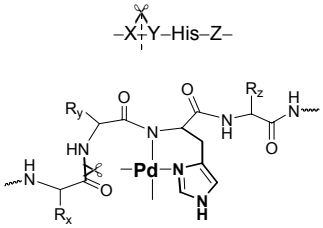
www.manaraa.com

Chart 1. Different cleavage selectivity of Pd(II) and Pt(II) complexes. Reproduced with permission from Inorganic Chemistry 2003, 42, 4036-45. Copyright 2003 American Chemical Society.

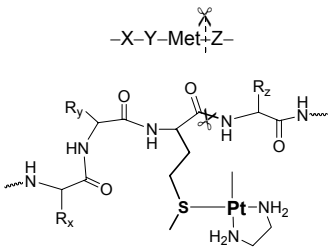
www.manaraa.com

$[Pd(H_2O)_4]^{2+}$ or *cis*-$[Pd(en)(H_2O)_2]^{2+}$





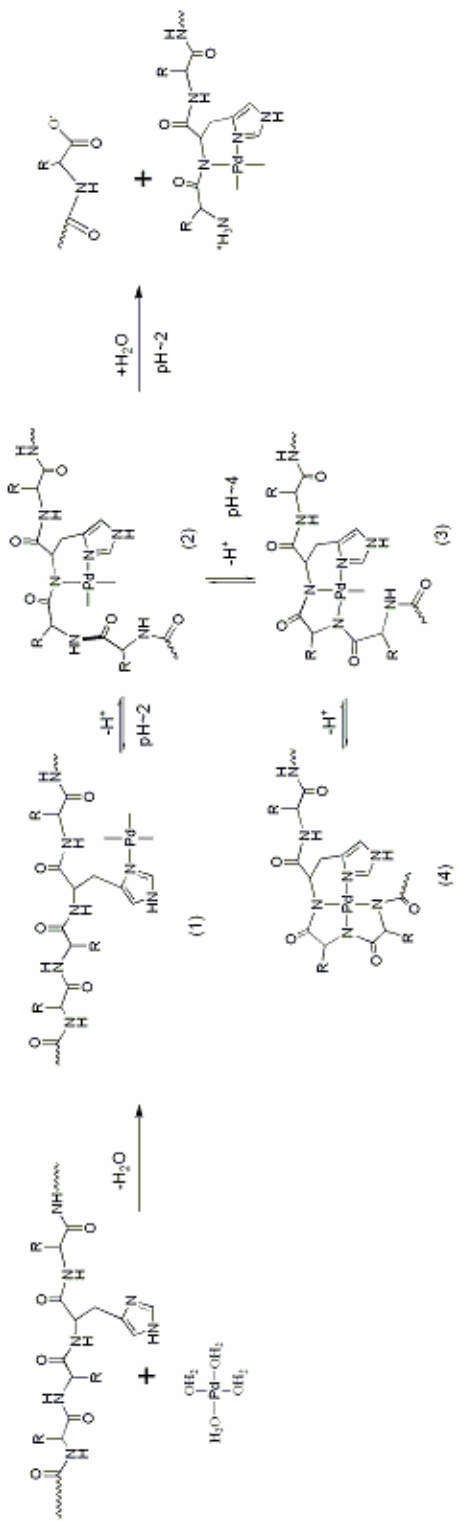
cis-$[Pt(en)(H_2O)_2]^{2+}$



www.manaraa.com

Figure 1. Cleavage of peptide bond by Pd(II) complexes. Coordination to Pd(II) of the methionine side chain (the anchor) is followed by coordination of the deprotonated nitrogen atom in the peptide backbone upstream from the anchor forming the hydrolytically active complex (complex 2). The pH of the solution determines the next step. If pH<2, hydrolytic cleavage occurs. If pH>2, the stepwise coordination of the peptide backbone takes place, with the formation of hydrolytically inactive complexes (complexes 3 and 4). The initial ligands and charges on the Pd(II) complexes are not shown for clarity.

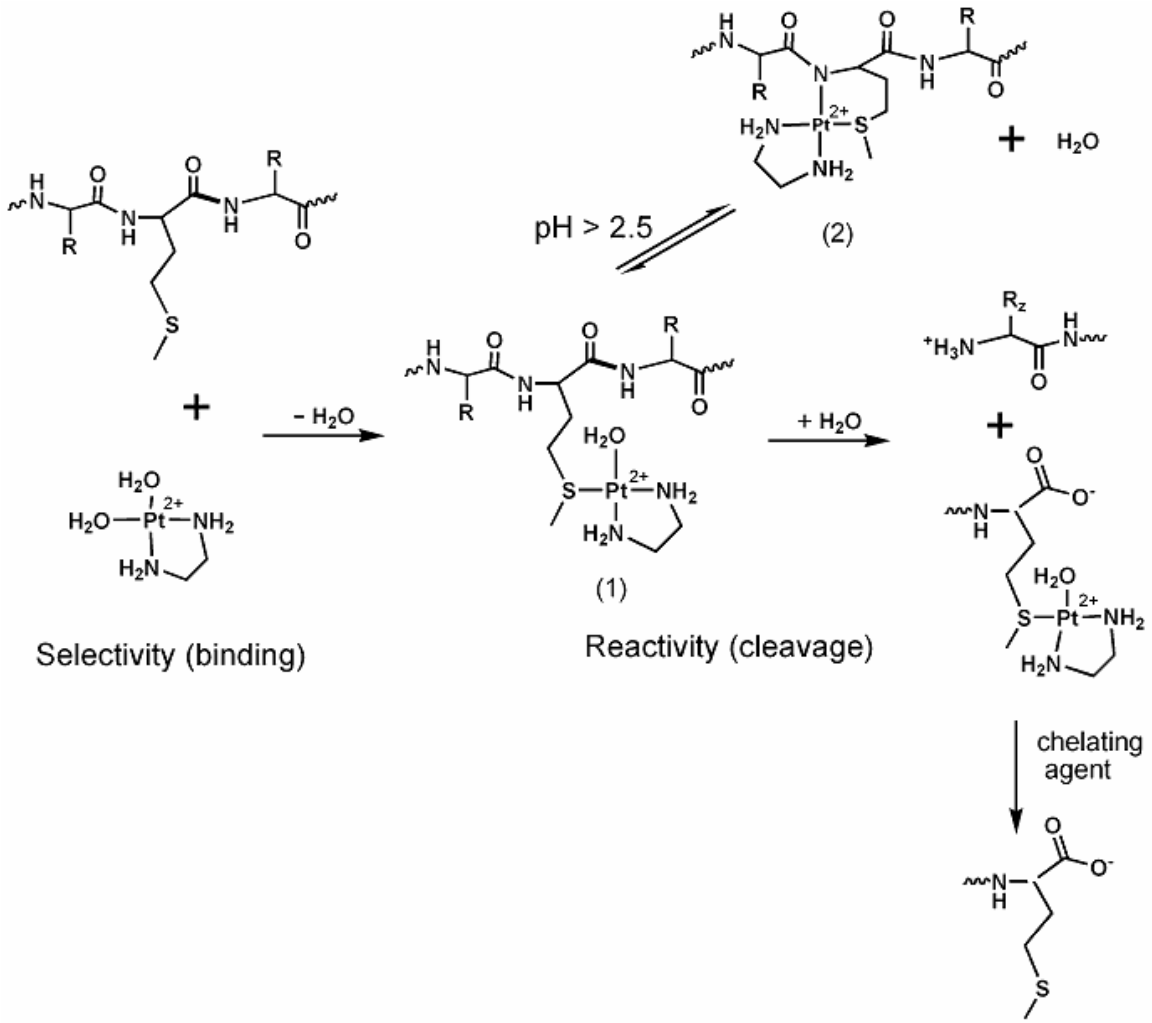
www.manaraa.com



www.manaraa.com

Figure 2. Cleavage of peptide bond by Pt(II) complexes. Coordination to Pt(II) of the methionine side chain (the anchor) results in the formation of the active complex (complex 1), followed by hydrolytic cleavage if pH<2.5. At a pH>2.5 the coordination of the deprotonated nitrogen in the peptide backbone downstream of the anchor to Pt(II) will result in the formation of an inactive complex (complex 2).

www.manaraa.com



www.manaraa.com

Chapter III. Platinum(II) complex as an artificial peptidase: selective cleavage of peptides and a protein by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ ion under ultraviolet and microwave irradiation

Laura-Mirela Dutcă, Kwang-Seuk Ko, Nicola L. Pohl, and Nenad M. Kostić

Department of Chemistry, Gilman Hall, Iowa State University, Ames, IA 50011

[1]A paper published in and reprinted from Inorganic Chemistry 2005, 44, 5141-51461

Abstract

Two synthetic peptides were completely cleaved by the *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex at pH 2.5 under thermal heating at 60°C in a selective way: only the amide bonds involving the carboxylic group of methionine residue, i.e., the Met-Z bonds (where the residue Z has a noncoordinating side chain) were hydrolyzed. Under irradiation at 300 nm, the rate constants for these cleavage reactions were approximately doubled, but side reactions occurred. Under microwave irradiation, the rate constants were increased two to three times at 60°C and ca. seven times at 100°C, and no side reactions were detected. Microwave irradiation similarly accelerated the complete and selective cleavage of Met-Z bonds in cytochrome *c* at 60°C in comparison with this cleavage under thermal heating, again without detected side reactions. The microwave-assisted cleavage of peptides and proteins by the platinum(II) reagent holds promise in proteomics and other biotechnological applications.

[1] Copyright 2005 American Chemical Society

www.manaraa.com

Introduction

Selective cleavage of peptides and proteins is essential in many bioanalytical and bioengeneering applications. Protein sequencing, peptide mapping,[1] folding studies,[2] protein semisynthesis,[3] and purification of fusion proteins all involve selective cleavage of peptide bonds.[4] The most desirable method of cleavage is hydrolysis of the amide group because the products of this reaction, namely amines and carboxylic acids, can be condensed into new products or otherwise chemically modified. Amide groups, however, are extremely unreactive toward hydrolysis; the half-life for peptide hydrolysis in the pH range from 4 to 8 is several hundred years.[5]

A small number of proteolytic enzymes and synthetic reagents are available, but they do not meet all current needs. Enzymes, such as trypsin and chymotrypsin, are very effective catalysts, but they have shortcomings. Their selectivity is almost fixed and very difficult to change, they become inactive in the presence of detergents, they are incompatible with organic solvents, they digest themselves as well as the intended substrate, and they contaminate the products of substrate cleavage. Sometimes these products (peptides) are so short as to be unsuitable for chemical ligation and other applications.

Chemical reagents are less effective than enzymes and have various disadvantages.[6] The common reagent cyanogen bromide cleaves at the C-terminus of methionine residues, but irreversibly converts these residues to serine lactone; it is volatile and toxic; it is applied in very large excess over the methionine residues present in the substrate; it requires 100% formic acid or 70% trifluoroacetic acid as a solvent; and

www.manaraa.com

it causes various side reactions.[6] Polymers having catalytic groups, such as carboxylate, aldehyde, and imidazolyl, show some promise as artificial peptidases.[7]

Transition-metal complexes have long been known to promote cleavage of peptide bonds, but their systematic study has only recently led to practical applications.[8-19] Currently, palladium(II) and platinum(II) complexes, in particular $[Pd(H_2O)_4]^{2+}$, *cis*-$[Pd(en)(H_2O)_2]^{2+}$, and *cis*-$[Pt(en)(H_2O)_2]^{2+}$, are the most effective inorganic reagents for protein cleavage.[14,20-26] Palladium(II) complexes have been studied in some detail. They spontaneously bind to methionine and histidine side chains and regioselectively promote hydrolytic cleavage of the second amide bond preceding this anchoring residue (in the direction of the amino terminus), that is, the X-Y bond in the X-Y-Met-Z and X-Y-His-Z sequences in which X, Y, and Z have noncoordinating side chains.[21-23,26] If Y is the proline residue, $[Pd(H_2O)_4]^{2+}$ can cleave the X-Pro bond at neutral pH.[23] The properties of the complexes can be adjusted, and they can cleave even in the presence of detergents[24] or in organic solvents.[27] Conjugates of Pd(II) complex and β-cyclodextrin can cleave selectively the X-Pro bond in the X-Pro-Ar sequence, where Ar is an aromatic residue, at neutral pH.[26] Kinetic and stereochemical evidence suggests that palladium(II) ion, as a Lewis acid, interacts with the carbonyl oxygen, thus polarizing the scissile amide group and facilitating nucleophilic attack of solvent water at the carbon atom.[23]

The study of platinum(II) complexes has only begun, and the results are interesting and unexpected.[24,25] Similar complexes of platinum(II) and palladium(II) generally undergo similar ligand-displacement reactions, but the former reacts much more slowly than the latter.[28] Surprisingly, the regioselectivity of *cis*-$[Pt(en)(H_2O)_2]^{2+}$ completely differs from that of *cis*-$[Pd(en)(H_2O)_2]^{2+}$, stated above. The platinum(II)

www.manaraa.com

complex binds only to methionine side chains and promotes hydrolytic cleavage of the first amide bond following this anchoring residue (in the direction of the carboxy terminus), that is, the Met-Z bonds, where Z is a noncoordinating residue.[24,25] This stark difference in regioselectivity can be attributed to the aforementioned difference between the metal ions. Because the platinum(II) complex is much more inert than its palladium(II) analog, the ethylenediamine ligand remains coordinated to the platinum(II) ion throughout the cleavage reaction, whereas this ligand is displaced by water at the palladium(II) ion early in the reaction.

The hydrolytic cleavage of proteins by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex is regioselective, but the reaction takes up to 24 h for completion, depending on the substrate and the reaction conditions. Although enzymes and conventional chemical agents require similar periods of time, we sought to make the platinum(II) reagent act faster. Even a twofold decrease in the reaction times would be a practical improvement. Slow chemical reactions can be accelerated by energizing the reactants, creating reactive intermediates, or stabilizing the products. Besides thermal heating, there is high pressure and irradiations with light, ultrasound, and microwaves. In most of these methods, energy in different forms is supplied to the reactants.[29]

Microwave (or dielectric) heating uses the ability of some compounds to transform electromagnetic energy into heat in situ. This is emerging as a new and promising method of accelerating chemical reactions.[30,31] The effect of temperature on the reaction rates is well known, but the effect of microwaves is not understood. Microwave irradiation can act through thermal effects or specific microwave effects. Thermal effects or dielectric heating can result from the interaction of polar molecules

www.manaraa.com

with the electromagnetic field. In liquids, only polar molecules selectively absorb the microwaves. Specific microwave effects are non-thermal, akin to the effects of the medium on the reaction mechanisms.[32]

In this study we explored the effects of ultraviolet light and demonstrated the effects of microwaves in accelerating selective cleavage of peptides and a protein by the *cis*-$[Pt(en)(H_2O)_2]^{2+}$ ion.

Experimental procedures

Chemicals. The complex *cis*-$[Pt(en)Cl_2]$ (en is ethylendiamine), piperidine, and α-cyano-4-hydroxycinnamic acid were obtained from Aldrich Chemical Co. The complex *cis*-$[Pt(en)(H_2O)_2]^{2+}$ was prepared by the published procedure as a perchlorate salt.[33,34] Its concentration was determined using the published absorptivity (extinction coefficient). Equine cytochrome *c* was obtained from Sigma Chemical Co. Trifluoroacetic acid was obtained from Alfa Aesar. Acetonitrile of HPLC grade was obtained from Fisher Scientific Co. All the Fmoc-amino acids, 1-[bis(dimethylamino)methylene]-1H-benzotriazoliumhexafluorophosphate(1-)3-oxide, 1-hydroxy-1,2,3-benzotriazole, FmocAla-Wang resin, and FmocGly-Wang resin, used in the synthesis of peptides, were purchased from Novabiochem.

The nonapeptide AcGly-Lys-Ala-Met-Ala-Ala-Pro-Arg-Gly (AcGKAMAAPRG) and the decapeptide AcAla-Lys-Tyr-Gly-Gly-Met-Ala-Ala-Arg-Ala (AcAKYGGMAARA) were synthesized by a standard manual Fmoc solid-phase procedure and purified by reverse-phase HPLC on a C18 preparative column, as described previously.[21,22] The purity, examined by analytical HPLC, was higher than

www.manaraa.com

99.5%. For the nonapeptide the found and calculated masses were, respectively, 901.84 and 901.07; for the decapeptide, 1036.57 and 1036.51 Da.

HPLC Separations. The components of the reaction mixtures were separated by a Hewlett Packard 1100 HPLC system containing an autosampler and a multiwavelength detector set to 215, 270, and 410 nm. Absorption at 215 nm is common to all peptides and proteins; absorption at 270 nm is due to aromatic residues and Pt(II) complex; and absorption at 410 nm is diagnostic of heme. The reverse-phase separations were done with an analytical Supelco Discovery C18 column (sized 250 x 4.6 mm, beads of 5 μm) and a preparative Vydac C18 column 218TP101522 (sized 250 x 22 mm, beads of 10 μm). The eluting solvent A was 0.10% (v/v) trifluoroacetic acid in water, and solvent B was 0.08% (v/v) trifluoroacetic acid in acetonitrile. For the reaction mixtures that involved the nonapeptide, AcGKAMAAPRG, in a typical run the percentage of solvent B in the eluent was kept at 0% for 5 min after the injection of the sample, and then raised gradually to 15% over a 35-min period. For the cleavage of the decapeptide, AcAKYGGMAARA, the method was the same, but the content of solvent B at 35 min was 45%. The flow rate was 1.00 mL/min in analytical runs and 10.0 mL/min in preparative runs. The size-exclusion separations were done with a Superdex peptide HR 10/30 column, having optimal separation range from 1000 to 7000 Da. The solvent was 0.10% (v/v) trifluoroacetic acid in water, and the flow rate was 0.50 mL/min.

Mass spectrometry. The MALDI-TOF experiments were done with a Bruker Proflex^{TM} instrument. The samples were prepared by a standard dried-droplet procedure: 1.0 μL of the solution of interest was mixed with 9.0 μL of a saturated solution of the matrix (α-cyano-4-hydroxycinnamic acid) in a 2:1 (v/v) mixture of water and

www.manaraa.com

acetonitrile. Each spectrum consisted of 100 scans. For the sake of clarity, molecular masses are reported only for the fragments free of the Pt(en) groups, although the Pt(en)-carrying fragments were also observed in the MALDI spectra. Bradykinin and cytochrome *c* were used as external standards. The measured molecular mass of a given fragment was compared with the value calculated by PAWS software, obtained from ProteoMetrics, LLC. An excellent agreement between the measurement and calculation conclusively identifies a peptide or relatively small protein.

Ultraviolet and microwave irradiations. The photochemical reactions were done in a Rayonet 100 reactor, which had 16 fluorescent tubes designated 3000 Å for the experiments at 300 nm, and eight fluorescent tubes designated 3500 Å for the experiments at 350 nm. The lamps have a bandwidth of approximately 25 nm on each side of the nominal emission maximum.

The experiments involving microwave irradiation were done with a CEM (Matthews, NC) Model Discover continuous-wave microwave oven delivering 300 W and allowing continuous cooling.[35]

Study of hydrolysis. In a typical experiment with ultraviolet irradiation, involving equimolar amounts of the Pt(II) reagent and the methionine residue in the substrate, 0.35 mL of a 60 mM solution of the nonapeptide, AcGKAMAAPRG, was mixed with 0.21 mL of a 100 mM stock solution of *cis*-$[Pt(en)(H_2O)_2]^{2+}$ and 1.44 mL of water. The final concentration of the peptide was 10.5 mM. The pH was adjusted by $HClO_4$ or NaOH. For a good comparison, the reaction mixture was divided in two 1.0-mL halves. One half was transferred to a quartz cuvette sized 10 x 10 x 40 mm with all walls transparent that had a rubber septum; thoroughly purged of air by bubbling with

www.manaraa.com

argon for 20 min; and irradiated by ultraviolet light for one day. The other half was kept in a 2.0-mL glass vial in the dark and heated in a dry bath (aluminum block). Samples of both halves were taken periodically; kept frozen, to quench the reaction; and analyzed by HPLC. After the reactions were completed, the pH remained within ±0.1 of the initial value. In the control experiments for possible background cleavage, the conditions were the same except that *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex was absent and the reaction was followed for much longer periods of time. The irradiated reaction mixtures had pH values of 2.0 and 2.5, and were kept at 40 and 60°C.

For the experiments involving microwave irradiation, the stock solutions were 5.00 mM in each substrate (the nonapeptide, the decapeptide, or cytochrome *c*). In a typical experiment, involving equimolar amounts of the Pt(II) reagent and the methionine residue in the peptide, the final volume of the reaction mixture was 5.00 mL, and the final peptide concentration was 1.00 mM. The reaction mixture contained 1.00 mL of the peptide solution, 50.0 μL of a 100 mM stock solution of *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex, and 3.95 mL of water. For the experiments with cytochrome *c*, the ratio of the protein to the Pt(II) reagent was 1:5. The pH was adjusted with $HClO_4$ or NaOH. The reaction mixture was divided in two parts; a 1.0-mL portion was kept in a dry bath, and a 4.0-mL portion was irradiated by microwaves. In the control experiments for possible background cleavage, the conditions were the same, except that *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex was absent and the reaction was followed for much longer periods of time.

The progress or absence of cleavage was monitored by size-exclusion chromatography in the case of the protein substrate and by reverse-phase HPLC in the case of the peptide substrates. In both cases the separated fragments were lyophilized to

www.manaraa.com

dryness, re-dissolved, and identified by MALDI-TOF mass spectrometry. This identification method is faster and more reliable than the sequencing of terminal residues, used in our earlier studies.[20]

Determination of the rate constants. Because the cleavage is very slow at the room temperature, at which HPLC was done, the species distribution in each chromatographic run matched that in the digest. The plots of the peak areas for the cleavage products versus time were fitted to the first-order rate law with the program SigmaPlot v. 5.0, obtained from SPSS Inc. All the kinetic plots have 5% error bars, reflecting the estimated inaccuracy in injecting the samples and integrating the peaks. Because the binding of the Pt(II) reagent to the methionine side chain is much faster than the subsequent intramolecular cleavage of the substrate, the fitting to the first-order rate law is justified. Each rate constant is the average of two consistent values, obtained by monitoring both fragments, products of the cleavage. The stated errors in the rate constants correspond to two standard deviations, i.e., confidence limit greater than 95.0%. These conservative error margins are our precaution against overstating small differences.

In experiments with irradiation at 300 nm, in which the product peaks increased for approximately three hours and then started to decrease, only the increasing part of the plot was fitted. This part corresponded to 85-90% of the cleavage reaction.

Results and discussion

Design of the photochemical experiments. The interaction of the platinum(II) ion with the scissile amide group probably involves dissociation of an aqua ligand. Therefore the ability of ultraviolet light to enhance ligand-substitution reactions at

www.manaraa.com

platinum(II) atom[36,37] may be relevant to the action of *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex in cleaving peptide bonds. We considered different irradiation wavelengths. The absorption spectrum of the complex shows a maximum at 256 nm, but we did not use the light of 254 nm lest it cause displacement of the ethylenediamine ligand. At 300 nm, phenylalanine, tyrosine, and tryptophan would be excited, since they absorb light around 280 nm. The presence of these residues might have caused side reactions that would have obscured the reaction of interest, the substrate cleavage by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex. For this reason, the methionine-containing nonapeptide, used in the experiments involving ultraviolet irradiation, lacked aromatic residues.

The cleavage of the nonapeptide by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex under ultraviolet irradiation. The equimolar mixture of the nonapeptide, AcGKAMAAPRG, and *cis*-$[Pt(en)(H_2O)_2]^{2+}$ ions at pH 2.50 and 60 °C was irradiated at 300 nm and analyzed by HPLC. Initially, two peaks were present: that containing the intact nonapeptide, eluting at 26.2 min, and that containing free *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex, eluting at 3.8 min. After four hours, two new peaks, eluting at 5.3 and 9.5 min, were the only ones present in the chromatogram. These two products of cleavage were identified by MALDI mass spectrometry. Very similar HPLC and MALDI results were obtained upon irradiation at 300 nm. Table 1 shows that cleavage by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex under ultraviolet irradiation occurs on the carboxy side of the methionine residue, that is, at the first amide bond "downstream" from the anchoring residue, as shown schematically below.

AcGly-Lys-Gly-Met-Ala-Ala-Pro-Arg-Ala

www.manaraa.com

The comparison of ultraviolet irradiation and thermal heating. Figure 1 shows a typical kinetic plot for the cleavage assisted by ultraviolet irradiation. Table 2 shows that the rate constants for the cleavage of the nonapeptide by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex were approximately two times higher under irradiation at 300 nm than under thermal heating, at all the pH values and temperatures examined. In other words, reaction time is approximately halved under irradiation at 300 nm.

Solutions irradiated at 300 nm, however, turned from light yellow to yellow and then to brown, while the thermally heated solutions stayed light yellow. Upon prolonged irradiation, HPLC peaks for the products decreased without new peaks emerging. Brown precipitate separated upon the centrifugation of the irradiated solutions. These symptoms of side reaction persisted when irradiation at 300 nm was done at 40°C and at a lower pH. These symptoms were less pronounced under irradiation at 350 nm, but, as Table 2 shows, the cleavage rate was only slightly higher than that under thermal heating. Although the photochemical method proved to be somewhat successful, we looked for other means of accelerating the cleavage reaction.

Effects of microwave irradiation on the reaction rates. Reactions of palladium compounds in homogeneous catalysis are markedly promoted by microwaves.[38,39] Very recently the microwave method began to be applied to platinum compounds,[40-42] but it has barely been used with peptides and proteins. Hydrolysis of these polyamides is much accelerated,[43,44] but the brutal acidic conditions used make the reaction nonselective and therefore good for protein sequencing but not for the production of discrete peptides in high yields. Selective cleavage at both carboxy-termini and amino-termini of aspartyl residues occurs in weakly acidic medium. The reaction time was three to six times shorter

www.manaraa.com

under microwaves than under conventional heating, but only 90% of the other peptide bonds stayed intact under the reaction conditions.[45] Microwave-enhanced cleavage by trypsin was recently used for protein mapping, but the reaction was incomplete.[46]

Because the thioether group is a fairly strong nucleophile for platinum(II) ion, displacement of an aqua ligand in *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex by the methionine side-chain occurs within minutes in our experiments. Because the carbonyl oxygen atom is a weak nucleophile, the interaction between the methionine-anchored platinum(II) ion and the scissile amide group is slow. This step and the subsequent external attack of the solvent water at the activated amide group may, in principle, be accelerated by microwaves. We set out to test this hypothesis.

The cleavage of oligopeptides by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex under microwave irradiation. Because microwave irradiation shortens reaction times and thus disfavors secondary reactions, we did the experiments at 100 as well as 60°C. To maximize the microwave energy imparted to the sample, we carried out experiments with continuous cooling. We continued to use the nonapeptide, AcGKAMAAPRG, so that we could compare various methods for promoting the cleavage. Since the microwave irradiation should not affect the aromatic residues, we also used the decapeptide, AcAKYGGMAARA. These two substrates gave consistent results. For example, an equimolar mixture of the nonapeptide and the Pt(II) reagent irradiated for 3 h at pH 2.5 and 60°C showed only two HPLC peaks, eluting at 5.3 and 9.5 min. Evidently, the cleavage was complete. As before, both fragments were identified by MALDI mass spectrometry. As Table 1, Table S1 in the Supporting Information, and the illustration

www.manaraa.com

below show, the selectivity under microwave irradiation is the same as that under ultraviolet light and conventional heating.

AcAla-Lys-Tyr-Gly-Gly-Met-Ala-Ala-Arg-Ala

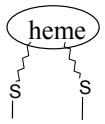
Nearly perfect match of measured and calculated molecular masses in Table 1 is evidence that the fragments retain their terminal amino and carboxylic groups. Even a slight chemical modification of the fragments would have affected their molecular masses, and MALDI-TOF spectra would have shown it. Evidently, the platinum(II) reagent cleaves the substrate by hydrolytic mechanism, as shown in Scheme 1.

Figure 2 shows that the half-life for cleavage of the nonapeptide is reduced to ca. 40 min at 60ºC. Table 3 shows that microwaves are generally two to three times more effective than thermal heating at 60°C in promoting cleavage by the Pt(II) reagent and six to seven times more effective at 100ºC. At this higher temperature the half-life for cleavage is only 3.5 min for the decapeptide and only 1.2 min for the nonapeptide. No side reactions were observed.

Encouraged by these results, we applied the new method to a protein.

The microwave-promoted cleavage of Equine cytochrome *c* by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex. Because the protein contains multiple residues capable of binding the reagent, we added five equivalents of the platinum(II) complex. The reaction at pH 2.5 and 60 °C was followed by size exclusion chromatography. Intact cytochrome *c* eluted at 16.5 min. After 12 h the intact protein was absent, and three fragments were identified by MALDI mass spectrometry. Table S2, in the Supporting Information, gives evidence for selective cleavage of Met65-Glu66 and Met80-Ile81 bonds, as shown schematically below.

www.manaraa.com



1 AcGDVEKGKKIFVQKCAQCHTVEKGGKHKTGP 30

31 NLHGLFGRKTGQAPGFTYTDANKNKGITWK 60

61 EETLMEYLENPKKYIPGTKMIFAGIKKKTE 90

91 REDLIAYLKKATNE 104

The growth of the peak corresponding to the fragment 81···104 at pH 2.5 and 60°C obeyed the first-order rate law under both thermal heating and microwave irradiation; see Figure 3. The respective rate constants were $(1.4 \pm 0.4) \times 10^{-3}$ and $(3.5 \pm 0.5) \times 10^{-3}$ min^{-1}. Microwave irradiation approximately doubled the cleavage rate, as in the experiments with oligopeptides. Because the cleavage time with our reagent was comparable to that with common proteolytic enzymes, we deemed the reaction sufficiently fast for practical work and did not do experiments at 100°C.

Conclusions and prospects

This study confirms that the complex *cis*-$[Pt(en)(H_2O)_2]^{2+}$ promotes selective, hydrolytic cleavage of peptide bonds involving the carboxylic group of methionine residue, i.e., the Met-Z bonds where Z has a noncoordinating side chain.[24,25] Irradiation at 300 nm increases the rate constant approximately two times, but this method is impractical because of side reactions. Microwave irradiation, however, increase the rate constant as much as seven times under conditions where side reactions are not observed. Two peptides and a protein were cleaved selectively and completely in a relatively short time. Because methionine residues are relatively rare in proteins, the products of cleavage

www.manaraa.com

are relatively long peptides, suitable for proteomics, semisynthesis, and other applications. Because cleavage of even large proteins gives relatively few peptides, our reagent may be useful in analyzing mixtures containing relatively many proteins. Even incomplete cleavage, achieved in yet shorter time, may be useful in proteomics applications.

Acknowledgement.

This work was supported by the National Science Foundation through grant CHE-0316868.

References

1. Hancock, W. S., Ed. *New Methods in Peptide Mapping for the Caracterization of Proteins*; CRC Press: Boca Raton, 1996.
2. Hubbard, S.; Beynon, R. J. *Proteolysis of Native Proteins as a Structural Probe*; Oxford University Press: New York, 2001.
3. Wallace, C. J. A. *Protein Engineering by Semisynthesis*; CRC Press: Boca Raton, FL, 2000.
4. *Applications of Chimeric Genes and Hybrid Proteins, Part A: Gene Expression and Protein Purification;* Thorner, J.; Emr, S. D.; Abelson, J. N.; Eds.; *Methods in Enzymology. 326; Academic Press: New York* 2000.
5. Radzicka, A.; Wolfenden, R. *J. Am. Chem. Soc.* 1996, *118*, 6105-6109.
6. Walker, J. M. *The Protein Protocols Handbook*; Humana Press: Totowa, NJ, 2002.
7. Suh, J. *Acc. Chem. Res.* 2003, *36*, 562-570.

www.manaraa.com

8. Allen, G. *Met. Ions Biol. Syst.* 2001, *38*, 197-212.

9. Buckingham, D. A. *Met. Ions Biol. Syst.* 2001, *38*, 43.

10. Datwyler, S. A.; Meares, C. F. *Met. Ions Biol. Syst.* 2001, *38*, 213-254.

11. Heyduk, T.; Baichoo, N.; Heyduk, E. *Met. Ions Biol. Syst.* 2001, *38*, 255.

12. Kito, M.; Urade, R. *Met. Ions Biol. Syst.* 2001, *38*, 187.

13. Komiyama, M. *Met. Ions Biol. Syst.* 2001, *38*, 25.

14. Milović, N. M.; Kostić, N. M. *Met. Ions Biol. Syst.* 2001, *38*, 145-186.

15. Polzin, G. M.; Burstyn, J. N. *Met. Ions Biol. Syst.* 2001, *38*, 108.

16. Kumar, C. V.; Buranaprapuk, A.; Cho, A.; Chaudhari, A. *Chem. Commun.* 2000, 597-598.

17. Rana, T. M.; Meares, C. F. *J. Am. Chem. Soc.* 1991, *113*, 1859-1861.

18. Hegg, E. L.; Burstyn, J. N. *Coord. Chem. Rev.* 1998, *173*, 133-165.

19. Kumar, C. V.; Buranaprapuk, A. *J. Am. Chem. Soc.* 1999, *121*, 4262-4270.

20. Zhu, L.; Qin, L.; Parac, T. N.; Kostić, N. M. *J. Am . Chem. Soc.* 1994, *116*, 5218-5224.

21. Milović, N. M.; Kostić, N. M. *Inorg. Chem.* 2002, *41*, 7053-7063.

22. Milović, N. M.; Kostić, N. M. *J. Am. Chem. Soc.* 2002, *124*, 4759-4769.

23. Milović, N. M.; Kostić, N. M. *J. Am. Chem. Soc.* 2003, *125*, 781-788.

24. Milović, N. M.; Dutcă, L.-M.; Kostić, N. M. *Chemistry Eur. J.* 2003, *9*, 5097-5106.

25. Milović, N. M.; Dutcă, L.-M.; Kostić, N. M. *Inorg. Chem.* 2003, *42*, 4036-4045.

26. Milović, N. M.; Badjić, J. D.; Kostić, N. M. *J. Am. Chem. Soc.* 2004, *126*, 696-697.

www.manaraa.com

27. Kaminskaia, N. V.; Johnson, T. W.; Kostić, N. M. *J. Am . Chem. Soc.* 1999, *121*, 8663-8664.

28. Cotton, F. A.; Wilkinson, G.; Bochmann, M.; Murillo, C. *Advanced Inorganic Chemistry, 6th Edition*, 1998.

29. Balzani, V.; Maestri, M. In *Photosensitization and Photocatalysis Using Inorganic and Organometallic Compounds*; Kalyanasundaram, K., Gratzel, M., Eds.; Kluwer Academic Publishers, 1993, pp 15-49.

30. Loupy, A., Ed. *Microwaves in Organic Synthesis*; Wiley-VCH: Weinheim, 2002.

31. Kappe, C. O. *Angew. Chem. Int. Ed.* 2004, *43*, 6250-6284.

32. Perreux, L. L., Andre In *Microwaves in Organic Synthesis*; Loupy, A., Ed.; Wiley-VCH: Weinheim, 2002, pp 61-114.

33. Heneghan, L. F.; Bailar, J. C., Jr. *J. Am. Chem. Soc.* 1953, *75*, 1840-1841.

34. Basolo, F.; Bailar, J. C., Jr.; Tarr, B. R. *J. Am. Chem. Soc.* 1950, *72*, 2433-2438.

35. Chen, J. J.; Deshpande, S. V. *Tetrahedron Lett.* 2003, *44*, 8873-8876.

36. Ford, P. C.; Hintze, R. E.; Petersen, J. D. In *Concepts of Inorganic Photochemistry*; Adamson, A. W., Fleischauer, P. D., Eds.; John Wiley and Sons, 1975.

37. Fry, H. C.; Deal, C.; Barr, E.; Cummings, S. D. *J. Photochem. Photobiol., A* 2002, *150*, 37-40.

38. Larhed, M.; Moberg, C.; Hallberg, A. *Acc. Chem. Res.* 2002, *35*, 717-727.

39. Olofsson, K.; Hallberg, A.; Larhed, M. In *Microwaves in Organic Synthesis*; Loupy, A., Ed.; Wiley VCH, 2002, pp 379-403.

www.manaraa.com

40. Adilia Januario Charmier, M.; Kukushkin, V. Y.; Pombeiro, A. J. L. *Dalton Transactions* 2003, 2540-2543.

41. Desai, B.; Danks, T. N.; Wagner, G. *Dalton Transactions* 2003, 2544-2549.

42. Desai, B.; Danks, T. N.; Wagner, G. *Dalton Transactions* 2004, 166-171.

43. Chiou, S. H.; Wang, K. T. *J. Chromatogr.* 1989, *491*, 424-431.

44. Zhong, H.; Zhang, Y.; Wen, Z.; Li, L. *Nat. Biotechnol.* 2004, *22*, 1291-1296.

45. Wu, C. Y.; Chen, S. T.; Chiou, S. H.; Wang, K. T. *J. Prot. Chem.* 1992, *11*, 45-50.

46. Pramanik, B. N.; Mirza, U. A.; Ing, Y. H.; Liu, Y.-H.; Bartner, P. L.; Weber, P. C.; Bose, A. K. *Prot. Sci.* 2002, *11*, 2676-2687.

www.manaraa.com

Table 1. Results of HPLC separation and MALDI mass spectroscopic identification of the fragments of AcGly-Lys-Ala-Met-Ala-Ala-Pro-Arg-Gly resulting from the cleavage by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex under ultraviolet irradiation

elution time (min)	molecular mass (D)		fragment
	observed	calculated	
5.3	448.66	448.56	1···4
9.5	472.94	471.54	5···9

www.manaraa.com

Table 2. Rate constants *k* for the cleavage of AcGly-Lys-Ala-Met-Ala-Ala-Pro-Arg-Gly by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex under thermal heating and under ultraviolet irradiation at the wavelengths shown

		$k/10^{-3}$ min^{-1}		
pH	T (°C)	thermal heating	300 nm	350 nm
	40	7±2	10±3	9±2
2.0	60	17±4	32±5	n.d.[a]
	40	5±1	11±2	4±1
2.5	60	13±2	20±4	n.d.[a]

[a] Not determined.

www.manaraa.com

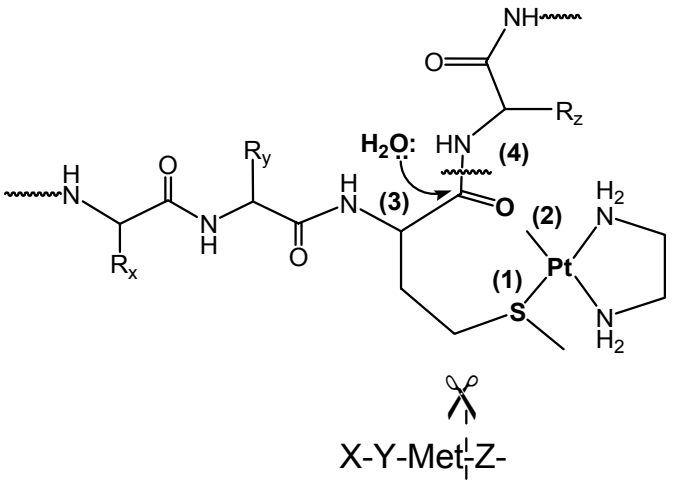
Table 3. Rate constants *k* for the cleavage of the nonapeptide AcGly-Lys-Ala-Met-Ala-Ala-Pro-Arg-Gly and the decapeptide AcAla-Lys-Tyr-Gly-Gly-Met-Ala-Ala-Arg-Ala by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex at pH 2.5 under thermal heating and microwave irradiation

T(°C)	$k/10^{-3}$ min^{-1} nonapeptide		decapeptide	
	thermal	microwave	thermal	microwave
60	6±1	17±3	2.5 ±0.5	7±1
100	80±11	580±35	33±3	200±25

www.manaraa.com

Scheme 1. Proteolytic selectivity of *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex and the four steps of cleavage of amide bonds by Pt(II) complexes: (1) binding of Pt(II) atom to the sulfur in the methionine side chain; (2) interaction of the Pt(II) atom with the neighboring amide group; (3) attack of the solvent water; and (4) hydrolysis of the amide group. The unspecified ligand on the Pt(II) atom is H_2O. Amino-acid residues X, Y, and Z have noncoordinating side chains.

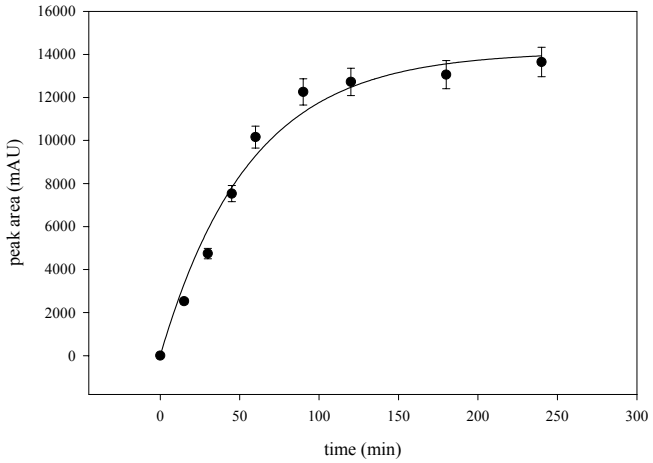
www.manaraa.com



www.manaraa.com

Figure 1. Kinetics of the cleavage of the nonapeptide AcGly-Lys-Ala-Met-Ala-Ala-Pro-Arg-Gly by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex under irradiation at 300 nm at pH 2.5 and 60°C. The appearance of the fragment 5···9 was followed by HPLC.

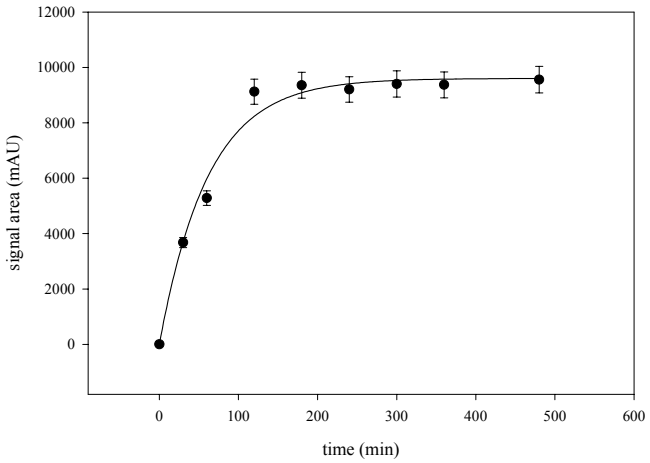
www.manaraa.com



www.manaraa.com

Figure 2. Kinetics of the cleavage of the nonapeptide AcGly-Lys-Ala-Met-Ala-Ala-Pro-Arg-Gly by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex under microwave irradiation, at pH 2.5 and 60°C. The appearance of the fragment 1···4 was followed by HPLC.

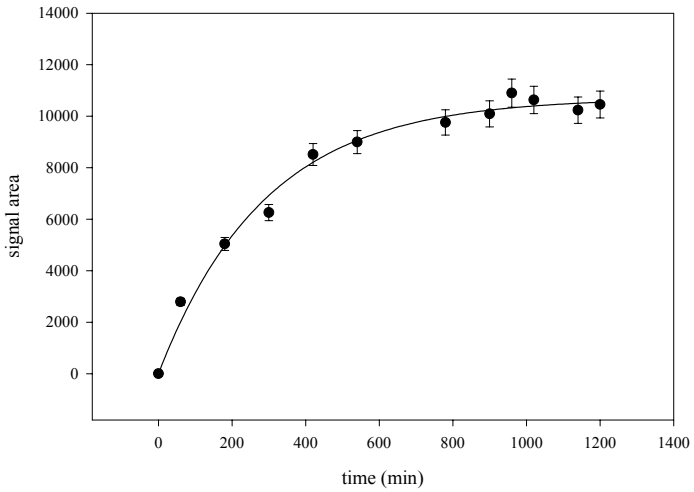
www.manaraa.com



www.manaraa.com

Figure 3. Kinetics of the cleavage of equine cytochrome *c* by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ complex under microwave irradiation at pH 2.5 and 60°C. The appearance of the fragment 81···104 was followed by HPLC.

www.manaraa.com



www.manaraa.com

Supporting information

Table S1. Results of HPLC separation and MALDI mass spectroscopic identification of fragments of AcAla-Lys-Tyr-Gly-Gly-Met-Ala-Ala-Arg-Ala (decapeptide) resulting from the cleavage by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ at pH 2.5 and 60°C under microwave irradiation.

elution time (min)	molecular mass (D)		fragment
	observed	calculated	
5.2	388.68	388.22	7···10
15.2	669.42	668.79	1···6

www.manaraa.com

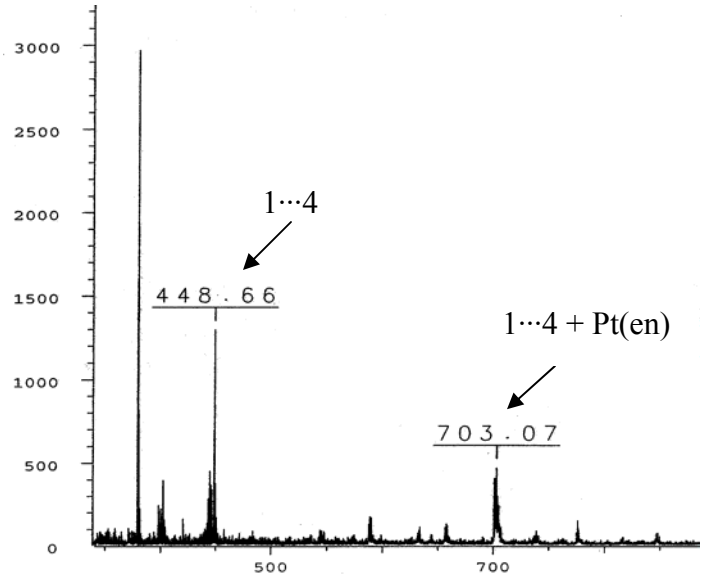
Table S2. Results of size-exclusion HPLC separation and MALDI mass spectroscopic identification of the fragments obtained from the cleavage of equine cytochrome *c* by 5 equiv of *cis*-$[Pt(en)(H_2O)_2]^{2+}$ at pH 2.5 and 60°C under microwave irradiation.

elution time (min)	molecular mass (D)		fragment
	observed	calculated	
18.5	7804.3	7802.6	1···65
23.8	2781.49	2780.3	81···104
28.4	1813.2	1811.1	66···80

www.manaraa.com

Figure S1. MALDI mass spectrum of the fragment 1···4 obtained in the cleavage of AcGly-Lys-Ala-Met-Ala-Ala-Pro-Arg-Gly (termed nonapeptide) by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ under irradiation at 300 nm. The peak at m/z =448.6 corresponds to the fragment 1···4, and the peak at 703.07 to the fragment 1···4 that carries a Pten group. The calculated molecular masses for the aforementioned species are 448.56 and 703.64. All the other peaks correspond to the matrix (α-cyano-4-hydroxycinnamic acid).

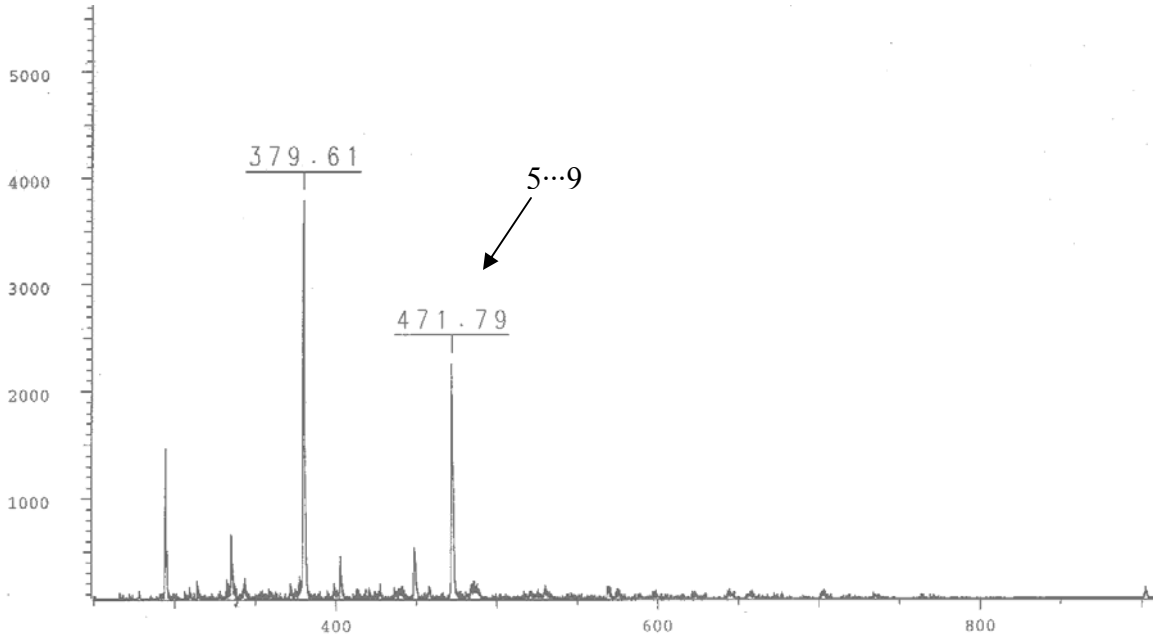
www.manaraa.com



www.manaraa.com

Figure S2. MALDI mass spectrum of the fragment 5···9 obtained in the cleavage of AcGly-Lys-Ala-Met-Ala-Ala-Pro-Arg-Gly (termed nonapeptide) by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ under microwave irradiation. The peak at m/z = 471.79 corresponds to the fragment 5···9, and the calculated molecular masses of the fragment is 471.54. All the other peaks correspond to the matrix (α-cyano-4-hydroxycinnamic acid).

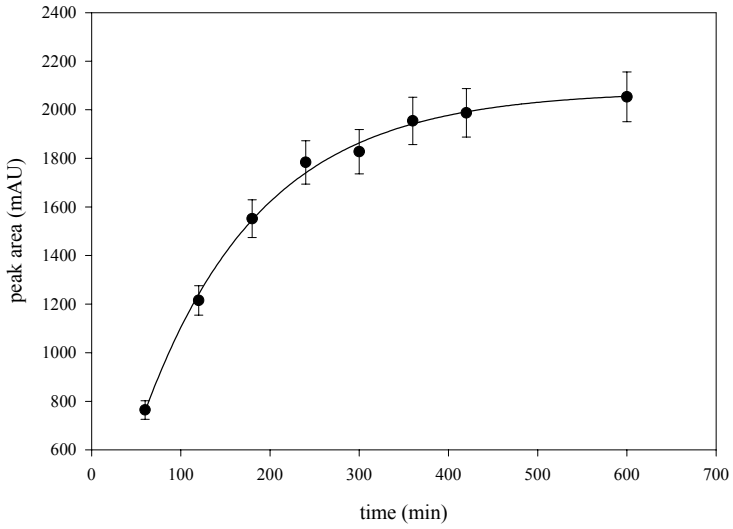
www.manaraa.com



www.manaraa.com

Figure S3. Kinetics of the cleavage of AcAla-Lys-Tyr-Gly-Gly-Met-Ala-Ala-Arg-Ala (termed decapeptide) by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ under irradiation at 300 nm, at pH 2.5 and 60°C. The HPLC signal for the fragment 7···10 grows in time

www.manaraa.com



www.manaraa.com

Chapter IV. Pt(II) complexes as selective protein cleavers. General summary and future directions

Pt(II) complexes can cleave selectively and efficiently Met-Z bonds in peptides and proteins in weakly acidic solution, at temperatures between 40 and 60°C, even in the presence of sodium dodecyl sulfate.[1-5] The same selectivity is observed when the reaction is performed under irradiation with ultraviolet or visible light, or microwave.[4]

Although Pt(II) and Pd(II) ions are chemically similar, their complexes cleave peptides with different hydrolytic selectivity.[2, 3] Pd(II) complexes cleave the second peptide bond upstream from the anchoring side-chain, histidine or methionine, the X-Y bond in the sequence X-Y-His(Met)-Z. This difference is explained by contrasting the modes in which each reagent binds to a specific side chain or to a side chain and the peptide backbone, and by their intrinsic reactivity. Pt(II) complexes selectively cleave the first amide bond downstream from methionine and are more suitable for producing longer peptide fragments, since the average abundance of methionine in proteins is about 2.2%. The combined abundance of methionine and histidine is about 5.5% , and obviously cleavage by Pd(II) would give rise to shorter fragments.

Hydrolytic cleavage of proteins by *cis*-$[Pt(en)(H_2O)_2]^{2+}$ (en is ethylendiamine) is accelerated under microwave irradiation. The reaction times were two to three times shorter under microwave irradiation, at 60°C and ca. seven times faster at 100°C, and no side reactions were detected for the peptides and the protein studied.[4]

The study of Pt(II) complexes as protein cleavage reagents is still seminal. It was shown that complexes containing ethylenediamine or 1,3-bis(methylthio)propane have

www.manaraa.com

the same selectivity, although cleavage is faster when ethylenediamine is present.[1-3,5] Nevertheless, there are still many unexplored direction. For example, until now, no studies were undertaken on the cleavage of peptides and proteins that contain cysteine residues, so more work is required, since cysteines are ubiquitous in peptides and proteins. One can envision that cleavage will take place at cysteine, but if it does not happen, there are reagents available to modify the sulfhydryl group, and overcome any interference from it.

There are numerous applications that one can envision for *cis*-$[Pt(en)(H_2O)_2]^{2+}$ as a cleavage reagent: generating large protein fragments that can be used in protein semysinthesis, as a reagent in peptide mapping by mass spectrometry, or other applications that require identification of fragments by mass spectrometry, for removal of expression tags, or to obtain structural information about proteins. The capacity of *cis*-$[Pt(en)(H_2O)_2]^{2+}$ to cleave even in the presence of detergents like sodium dodecyl sulfate could make it a useful tool in the study of insoluble membrane peptide and proteins. The wide array of applications that seem suitable for *cis*-$[Pt(en)(H_2O)_2]^{2+}$ make it a promising reagent for the study of biological systems.

References

1. Burgeson, I. E.; Kostic, N. M., *Inorg. Chem.* **1991,** 30, (23), 4299-4305.
2. Milovic, N. M.; Dutca, L. M.; Kostic, N. M., *Chemistry* **2003,** 9, (20), 5097-106.
3. Milovic, N. M.; Dutca, L. M.; Kostic, N. M., *Inorg. Chem.* **2003,** 42, (13), 4036-45.
4. Dutca, L. M.; Ko, K. S.; Pohl, N. L.; Kostic, N. M., *Inorg. Chem.* **2005,** 44, (14), 5141-6.

www.manaraa.com

5. Stoffregen, S. A.; Griffin, A. K.; Kostic, N. M., *Inorg. Chem.* **2005,** 44, (24), 8899-907.

www.manaraa.com

Chapter V. RNA-protein interactions in the 30S ribosomal subunit

Introduction

Probing RNA structures with different chemical reagents is a powerful tool in the study of RNA-protein interactions. One of the systems for which probing with different chemical reagents was applied is the small (30S) subunit of the ribosome. In this chapter, the 30S subunit and some approaches for its study involving chemical reagents, as base-specific footprinting and hydroxyl directed probing, are presented.

RNA-protein interactions play very important roles in many central cellular functions, such as replication, transcription and translation. RNA and proteins bind to form ribonucleoprotein particles (RNPs) such as the ribosome or spliceosome. The assembly of complex RNPs involves ordered binding of multiple proteins and is an intricate process. In these RNPs, RNA is often a ribozyme, or essential for catalysis.[1-4] Many RNPs, like the ribosome, contain large RNA molecules which can adopt numerous conformations. Generally, proteins help the RNA adopt the conformation necessary to form a functional RNP, and they are sometime referred to as "RNA glue".[5] Thus, while not necessarily catalytic, proteins are essential for active site organization.

Often, binding of protein to RNA involves major conformational changes in one or both the RNA and protein. Dynamically disordered parts of either one of the binding partners can adopt a defined conformation in the complex. The binding process that involves conformational changes in one or both of the partners is termed induced fit.[6,7] An important goal of structural biology is to understand the mechanistic and energetic consequences of the induced fit mechanism that is so ubiquitous. These conformational

www.manaraa.com

changes are thought to be important for affinity and specificity in RNA-protein interactions, and sometime for the function of the intact complex. For a better understanding of such conformational changes it is necessary to study RNA-protein interactions, by different approaches, in well characterized model systems. The small subunit of the prokaryotic ribosome is such a suitable model system.

30S ribosomal subunit - a model system for the study of RNA-protein interactions

The prokaryotic ribosome. The ribosome translates the message encoded in the nucleotide sequence of the messenger RNA into the amino acid sequence of a protein. It provides the structural framework for the decoding process and contains the catalytic center responsible for the formation of peptide bond. Ribosomes from different types of cells have the same basic structure but vary in size. All prokaryotic ribosomes (70S) are composed from two asymmetric RNPs, a large 50S subunit and a small 30S subunit. The nomenclature of the subunits of the ribosome derives from the experiments in which ribosomes were isolated from cell lysates by utracentrifugation.[8] The RNPs were named according to their sedimentation characteristics during centrifugation, which are determined by the molecular size and geometrical shape of the complex.[8] Svedberg units (S), the units of measurement used for ultracentrifugation, are not additive. The Svedberg values of the two ribosomal subunits do not add up to that of the entire ribosome, due to the loss of surface area when the two subunits are bound. The size of the prokaryotic ribosome is 2-2.5 nm and its approximate mass is 2.6-2.8 kDa.[9] The crystal structure of the ribosome from different prokaryotic organisms has been determined.[10-13] The

www.manaraa.com

ribosome is a very complex and massive macromolecular machine and often, for a better understanding of the RNA-protein interactions that take place, its subunits are studied separately. The small 30S subunit, our model system for the study of RNA-protein interactions and assembly of RNPs, binds messenger RNA, initiation factors, the large subunit 50S and participates in transfer RNA selection, during translation.

30S subunit - structure and composition. The 16S ribosomal RNA (rRNA) and 21 ribosomal proteins (r-proteins) (S1-S21) interact to form the 30S subunit. The proteins have the letter "S" in their name to indicate that they are components of the small subunit. The 30S subunit has a molecular weight of approximately 0.85 kDa and one-third of the mass consists of r-proteins, while the remaining mass the rRNA.[9] The crystal structure of the ribosome from *E. coli* (Figure 1) was determined recently.[13]

16S rRNA. The primary sequence of 16S rRNA from *E. coli* was determined in the late seventies,[14] and the secondary structure was elucidated by phylogenetic comparison and footprinting studies.[15] The sequence of 16S rRNA is highly conserved among all organisms.[16] 16S rRNA is composed from four domains which assemble with the corresponding r-proteins into different parts of the 30S subunit (Figure 2a). The four domains and their corresponding parts of the 30S subunit are: 5' domain of 16S rRNA which forms the body of the 30S subunit, the central domain which folds in the platform, the 3' major domain that forms the head and the 3' minor domain which mainly forms the penultimate stem[13, 15] (see Figure 2a and b). Each of the aforementioned domains of 16S rRNA can assemble independently of the others, in the presence of the correct r-proteins.[17, 18]

www.manaraa.com

Small subunit r-proteins. The *E. coli* r-proteins are small in size up to 250 amino acids, with the exception of S1 which has about 550 amino acids. In general, r-proteins are very basic, with an average pI of 10, which is not surprising since they interact with RNA which is a negatively charged molecule.[19, 20] From a structural point of view, they typically have one or more globular domain(s) and they often contain extended internal loops or long N- or C-terminal extensions.[19] Usually the loops and the extensions are very rich in basic amino acids such as arginine and lysine and they are closely associated with the rRNA.[13, 19] The r-proteins are predisposed to interact with rRNA through salt-bridges, through the positively charged residues, which suggests that shape and complementarity rather than sequence–specific interactions define the rRNA-r-protein interactions.

The small subunit r-proteins are classified from assembly point of view in three classes: primary (S4, S7, S8, S15, S17 and S20), secondary (S5, S6, S9, S11, S12, S13, S16, S18 and S19) and tertiary (S2, S3, S10, S14 and S21) (Figure 2c). Primary binding proteins bind directly and independently to16S rRNA, secondary binding proteins require the prior binding of at least one primary protein and tertiary binding proteins involve the initial assembly of at least one primary and one secondary binder. More details on the assembly of the 30S subunits and the classification of small subunit r-proteins are given in a later section. The primary binding r-proteins are prominently used in our studies and they are presented in more detail below.

Ribosomal proteins S4, S7, S8, S15, S17 and S20

There are six r-proteins which interact independently and directly with 16S rRNA.[21] Each of these r-proteins is thought to be important for direct rRNA

www.manaraa.com

conformational rearrangements and they enable other r-proteins to assemble in latter assembly events. Three of them (S4, S17 and S20) bind the 5' domain and help in the organization of the body of the 30S subunit.[13,19,22,23] The remaining r-proteins S8 and S15 bind to the central domain to organize the platform, while S7 binds to the 3' major domain to form the head of the small ribosomal subunit.[13,19,24] The 3' minor domain forms the penultimate stem and helix 45, and has only one primary binding r-protein that interacts with it, S20, which binds also to the 5' domain[13, 19, 22] (Figure 2 and Figure 3). While all primary r-proteins are likely important for the assembly cascade, S4 and S7 are considered the initiators for the assembly of the 30S subunit.[25]

One of the few r-proteins that interacts with two different domains of 16S rRNA is S20, as it was shown by footprinting[22] and the crystal structure of the 30S subunit. Initially there was some discrepancy in the placement of S20. Neutron diffraction mapping placed it in the head of the 30S subunit,[26] while immunoelectron microscopy and footprinting studies located it at the bottom of the body.[22] Directed hydroxyl radical probing also positioned it at the bottom of the 30S subunit.[27] Crystal structure of the 30S subunit showed that S20 is a three-helix bundle located at the bottom of the 30S subunit, and it binds several helices from the 5' domain (body) and the 3' minor domain (penultimate stem).[13, 19] Its structure has not been determined in the free form, when not bound to the 16S rRNA. Based on the model of 5' to 3' assembly[28] and the positioning of helix 44 across the body in the small subunit,[13] it is easy to speculate that S20 may interact with these domains differently along the assembly path.

The r-protein S17 binds both the 5' domain and the central domain, as it was observed in the crystal structure.[13, 19] Its location is at the interface formed between the

www.manaraa.com

top right of the body as viewed from the interface side that includes helices 7 and 11, and the three-way junction H20/H21/H22 of the platform (Figure 2). The structure determined in isolation by NMR showed that the core of *Bacillus stearothermophilus* (*B. st.*) S17 consists of a β-barrel with the oligonucleotide/oligosaccharide-binding fold, a motif common for a few classes of RNA binding proteins.[29] Beside the β-barrel, long disordered loops were also observed in the free form. For both organisms, *E.coli* and *B. st.*, the core of S17 is involved in extensive contacts with H7 and H11 in the 5' domain, in the 30S subunit and stabilizes the sharp bend at the H7-H11 junction. The chemical footprints observed for S17 are present mainly in helix 11 of the 5' domain,[22] and no footprints were observed in the central domain. It is possible that in the minimal complex S17/16S rRNA the central domain is not well organized, thus the central domain footprints are not observed. Majority of the residues from *E. coli* S17 that contact the central domain are in the loop 25-40,[13,19] which is presumably quite dynamic while the central domain is not assembled, as in the crystal structure of S17 in isolation.[29] The complex is probably not stable enough to be seen via footprinting or the interaction with the central domain may occur latter in the assembly.

S4 nucleates assembly of the 5'domain and it is the main protein stabilizing the back of the shoulder of the 30S subunit.[13,19,23] The r-protein S4 binds the five-way junction formed by H3, H4, H16, H17 and H18 (Figure 2a and b, and Figure 3). The structure of S4 in isolation without the first 41 amino acid residues was determined by both crystallography[30] and NMR[31], and both structures are very similar to the corresponding fragment of the r-protein in the 30S subunit.[13,19] But, since a significant

www.manaraa.com

part of S4 is missing, it is not known how that piece behaves at binding to 16S rRNA. More details on the interaction of S4 with 16S rRNA are presented in a later section.

The r-protein S15 binds the three-way junction between H20, H21 and H22[13, 19] (Figure2a and b, Figure 3). The protein has a simple four-helix bundle structure similar to S20, and its structure was determined in isolation by both NMR and X-ray crystallography for different organisms.[32] The structures of S15 in free form, bound to a fragment of RNA, as part of the low-resolution structure of the central domain of 30S subunit, and in complex with an RNA fragment, S6, and S18, are quite similar to the one of the r-protein in the 30S subunit.[13, 19, 33] The footprints specific for S15 are localized in helices 22 and 23[24] and they are very consistent with the S15 binding site. Thus, S15 interactions are restricted to the central domain of 16S rRNA.

S8 is located near the center of the back of the body in the structure of the 30S subunit and it may play a critical role in orienting the platform (central domain) relative to the body (5' domain) of the 30S subunit (Figure 3). S8 binds near the H20/H21/H22 three-way junction making extensive interactions with H21 and H25.[13] Hydroxyl radical footprinting and probing data are in good agreement with the crystal structure of the small ribosomal subunit. The N-terminus of the r-protein packs against the helix H25, thus helping the folding of the central domain.[13] The crystal structure of S8 from *B.st* and *Thermus thermophilus* (*T. thermophilus*) was also determined in isolation and it is very similar to the one from the 30 S subunit of *T. thermophilus*.[19]

The footprints specific for S8 are present throughout the central domain and in the 5' domain.[24] More precisely S8 footprints in the 530 loop, 570 region, helices 20, 21 and

www.manaraa.com

23, and also the 820 and 860 regions.[24] Obviously, S8 organizes more extended regions, not only the S8-16S rRNA direct surface contact.

S7 nucleates the assembly of the head, by binding to two multiple-stem junctions H28/H29/H43 and H29/H30/H41/H42 from the 3' domain (Figure 2a and Figure 3). The structure of S7 from *B. st.* and *T. thermophilus* was determined in isolation by crystallography, and a triangular-shaped helical domain with a highly conserved β-hairpin extension was observed. In the 30S subunit the very basic N-terminus of S7 was disordered in isolation but adopts a clear conformation when bound to the rRNA.[13] In the *E. coli* S7 (in the K strains) 20 additional residues are present in the C-terminus. The structure of S7 from *T. thermophilus* in isolation is almost identical with the one in the 30S subunit, except the orientation of the β-hairpin. The orientation of the hairpin with respect to the helical domain is different in the structure of the small subunit from either of the isolated structures.

Binding of S7 affects the reactivity of an extensive region in the lower half of the 3′ major domain.[34] The number of footprints specific for S7 is quite high and they are caused by a combination of direct RNA-protein contacts and S7 induced packing of minor groves.[13,19,34] Many of the footprints are present in regions where there are direct S7-16S rRNA contacts (the two multiple stem junctions mentioned above), but there are some in areas situated at some distance from the binding interface (region 980, and loops 1330 and 1360).[13,34]

The structure of some of the primary r-proteins (S4, S7, S8, S15 and S17) in free form was also resolved by NMR, X-ray crystallography, or both. Identifying changes that occur during the assembly of 30S subunit, in the rRNA and the r-proteins is the next

www.manaraa.com

challenge in understanding the assembly of the small subunit of the ribosome. The crystal structures revealed the structure of the final product of the assembly, or of the separate components, without clarifying how the ribosome assembles from its components. Other approaches have to be used to decipher the principles of RNP assembly and how RNA-protein interactions help in this process. The structure of the final product will help in the interpretation of the results obtained by other approaches.

Assembly of the 30S subunit. One of the characteristics that make 30S subunit such an attractive model system for studies of RNA-protein interactions is the fact that it assembles *in vitro* from its components.[21] Furthermore, it reconstitutes not only when using a mixture of r-proteins extracted from natural 30S subunits,[21] but also when individually purified proteins are used.[35] Functional 30S subunits were also reconstituted using recombinant proteins.[36,37] Each of the ribosomal proteins (S2-S21) was cloned, overexpressed, purified and assembled with natural 16S rRNA to form functional 30S subunits.[36,37] Extensive biochemical and genetic manipulation can be performed with this system to understand the functional role of any particular r-protein of interest and the nature of protein-RNA interactions that constitute the 30S subunit. The recombinant system makes possible some of the experiments that will be presented in this thesis.

The *in vitro* reconstitution system made possible the determination of the 30S subunit assembly map[38, 39] (Figure 2c) which depicts the protein dependencies for binding 16S rRNA during assembly. Assembly of the 16S rRNA into its functional conformation from 30S subunit is orchestrated by sequential binding of r-proteins. The r-proteins have been categorized into the three assembly classes (primary, secondary and tertiary)[35,38,39] mentioned earlier, as it is shown in the assembly map (Figure 2c).

www.manaraa.com

The assembly of the 30S subunit *in vitro* is highly temperature dependent. Reconstitution experiments performed at different temperatures revealed two intermediates with virtually the same composition but different sedimentation coefficients, RI (Reconstitution Intermediate) and RI*. RI is formed at low temperature (0-15°C) and the activated intermediate RI* is obtained from RI by heat activation.[40-42] A large conformational change is responsible for the conversion of RI into RI*. The r-proteins present in the two assembly intermediates are a subset of the primary and secondary r-proteins.[40-42] After activation, the remaining r-proteins are able to bind to the RI* and form functional 30S subunits. *In vivo*, the strains that have 30S subunit assembly defects are cold-sensitive and from these defective strains assembly intermediates similar to those observed *in vitro* were also isolated.[40,43,44] These observations suggest that the intermediates observed *in vitro* are true representatives of the assembly pathway and not merely experimental artifacts.

RNA-protein interactions in the 30S subunit

30S ribosomal subunit has 22 components which interact in an ordered fashion to form the macromolecular complex. The complexity and size of this system makes its study quite complicated. The crystal structure of the ribosome and of the separate units gave us a clear picture of the final product but it does not reveal how all these components come together to form the RNPs. Other methods that analyze RNA- protein interactions play an important role in the elucidation of the mechanism of assembly of the 30S subunit.

An intricate network of intramolecular and intermolecular interactions is involved in the process of 30S subunit self-assembly. There are obviously three different classes of

www.manaraa.com

interactions present in the small subunit: RNA-RNA, RNA-protein and protein-protein interactions. Since in our studies we are exploring mainly RNA-protein interactions methods suitable for their investigation are presented. Often these methods are used with the adequate adaptation, for the study of other types of interactions. The main biochemical approaches used in the identification of the interaction sites of r-proteins with 16S rRNA and to understand the effects of the RNA-protein interaction on the structure of the 16S rRNA are footprinting, cross-linking and binding assays. The first question that needs to be answered, especially when studying a large RNA molecule like 16S rRNA is to what part of the rRNA is the r-protein binding. All three of the aforementioned methods can give an answer to this question. Cross-linking gives information describing contacts between different rRNA regions (intra-RNA cross-linking) and contacts between16S rRNA and individual ribosomal protein (RNA-protein cross-linking). Since cross-linking requires connection of two or more partners through covalent bonds, the region(s) of 16S rRNA to which r-proteins bind can be determined. Reagents like bis(2-chloroethyl)methylamine[45-47] and 1-ethyl-3-(3-diethylaminopropyl)carbodiimide[48-50] were used to study the interaction of r-proteins with 16S rRNA.

Binding studies using methods such as filter binding assay,[51-54] gel mobility shift assay,[55] and sucrose gradient assay[56] give information on the strength of the RNA-protein interaction and reveal minimal binding sites. Footprinting with a few different types of probes can also reveal the binding interface.[57,58] Establishing the binding site and the strength of the interaction is the first step in the study of RNA-protein interaction. Another important question is what happens to the binding partners during the

www.manaraa.com

interaction. Is their conformation changing after binding, if there is a change is it localized to the contact surface or it does propagates at some distance? Cross-linking and binding studies can give some answers to these questions, but footprinting is probably the most powerful biochemical method to study conformational changes that take place in RNA. Footprinting combined with primer extension make possible the examination of the rearrangements that take place in 16S rRNA during interaction with r-proteins, by determining the changes in reactivity of almost each nucleotide.[57,58] The variations in reactivity are attributed to differences in the accessibility of nucleotides in the probed molecule, from which conformational changes can be determined.[57,58] Base specific probes modify nucleotides, and they are used to identify the nucleotides that undergo changes at the binding of r-proteins. Hydroxyl radicals cleave the RNA backbone, and reveal regions of 16S rRNA that become protected at the binding of r-proteins. A combination of the two types of probes identifies nucleotides which are involved directly or indirectly in binding. Besides giving information on the alteration in reactivity around the binding site, footprinting will reveal long distance conformational changes that take place as a result of the interaction.[57,58]

Site-directed hydroxyl radical probing of RNA by using Fe(II) tethered to unique positions on individual proteins is a different approach used to gain information about the three-dimensional rRNA environment around the tethered Fe(II) probe.[59] Primer extension is used to map the changes in reactivity and the cleavage sites. Since footprinting and primer extension are used in our studies, a more detailed description is presented in one of the next sections.

www.manaraa.com

Probing of RNA conformation is one of the examples where chemical reagents are preferred over enzymes. In the 1960s and 1970s enzymatic digestion was used to study the interaction between 16S rRNA and r-proteins. Enzymatic digestion takes advantage of the fact that the susceptibility of RNA to digestion depends on its conformation, or use nucleases that prefer single-stranded RNA to double-helical RNA. Nevertheless, chemical reagents became the favorite probes for the study of conformational changes in RNA over enzymes, and size was the main advantage that they had. Enzymes are much larger than the structural details they were used to study, and scientists were concerned that reactivities might depend on more than just the local conformation. Beside that, there was not enough information on the selectivity of these enzymes, regarding RNA structures other than the A-form.[60]

Dynamics of the r-protein 16S rRNA interaction

Co-transcriptional assembly. Kinetic footprinting studies revealed the sequential and cooperative nature of assembly.[28] The differential change in the reactivity of nucleotides during assembly at different temperature was monitored. This study made possible a different classification of r-proteins into different kinetic groups based on the temperature and thus the order in which the r-proteins bind to the growing RNP. In the earlier footprinting studies the footprints specific for a certain protein were identified, and the change in reactivity of nucleotides was attributed to binding of specific proteins. The r-proteins were classified in early binders (S4, S6, S11, S15, S16, S17, S18, S20), mid binders (S7, S8, S9, S13, S19), mid-late (S5 and S12) and late binders (S2, S3, S10, S14, S21) based on the changes in their footprints during assembly.[28] Even though each protein belongs to one kinetic class, the footprints attributed to it can belong to different

www.manaraa.com

kinetic classes. R-proteins that bind to the 5' and central domains fall into the early binding class and proteins that bind to the 3' major domain fall into the mid and late assembly groups. These results suggest that the 5' and central domains start assembling prior to the 3' domain. Three of the six proteins that interact with the 5' domain are primary binders and there are no tertiary binding proteins in this region (Figure 2c). At the other end is the 3' major domain where S7, a mid binder, is the only primary protein present, that nucleates folding and five of the eight proteins that bind to this domain are tertiary proteins (Figure 2c). Since the r-proteins that bind to the 5' domain bind the fastest and with high affinity, it seems plausible that the 5' region might be the first to fold as soon as transcribed and that 3' major domain assembles last.[28]

Conformational changes in 16S rRNA during assembly. A detailed investigation of the changes in the conformation of 16S rRNA during the assembly was performed, by chemical modification and primer extension analysis of each assembly intermediate.[61,62] The 16S rRNA undergoes major conformational changes induced by the binding of r-proteins to form RI, while rearrangements induced by concerted action of temperature and r-proteins take place in the heat activation step from RI to RI*. In the last step which is not temperature dependent, binding of the remaining r-proteins will produce a functional 30S subunit. During the early stages of assembly majority of the changes are present in the 5' domain while as assembly proceeds, the area where the changes are observed shifts towards the 3' end of 16S rRNA. These observations, similarly to the ones from the kinetic footprinting studies confirm a polar nature of the assembly, which might be reflective of co-transcriptional assembly *in vivo*. The roles of r-proteins at

www.manaraa.com

different stages of assembly were also dissected and the most important proteins for each step were proposed.[61,62]

Assembly landscape. Another approach used kinetics of binding of r-proteins to follow the assembly of 20 r-proteins with 16S rRNA to form the 30S subunit by pulse–chase quantitative mass spectrometry.[63] The protein binding rates to 16S rRNA or 16S rRNA containing RNPs at a range of temperatures were determined. A very complex assembly process in which different pathways are available that converge in the final point, formation of a functional 30S subunit, was revealed. An assembly landscape for the formation of the functional subunit was determined and based on it, the binding of each r-protein further stabilizes the native 30S conformation, until all assembly pathways converge at this state.[63]

The aforementioned studies of the dynamics of the rRNA-r-protein interactions during assembly analyzed the global changes of the 16S rRNA in the presence of all or many r-proteins. The interaction between the 16S rRNA and each of the individual r-proteins and the importance of each r-protein in the assembly process was not dissected. When all the r-proteins are present the concerted changes in the conformation of the 16S rRNA might obscure the contribution of each protein. Other studies, like the ones that we will present in this thesis can reveal more on the role of the r-proteins in the assembly of the 30S subunit.

A temperature-dependent conformational rearrangement in the S4/16S rRNA complex. In the eighties two binding sites for S4 were determined by different methods.[23,53,64] Chemical footprinting by the Noller lab, showed that a number of bases in the 5' domain were protected by bound S4, bases confined mainly to the

www.manaraa.com

H4/H16/H17/H18 helical junction,[23] which was termed the ”S4 junction”. This result was also supported by direct RNA-protein cross-linking[64] to one of the helices and the earlier observations of nuclease protection studies. The Draper lab studied the interaction of 16S rRNA with S4 by determining the binding constants with different synthetic subfragments of 16S rRNA.[53] Omission of regions of the rRNA containing two of the five implicated helical elements did not influence the specific binding of S4 to the RNA, indicating a smaller binding site for S4 then the one determined by footprinting.[53, 65] One difference between the footprinting and the binding experiments was that they were performed at different temperatures, at 42°C[23] and 0°C respectively.[53, 65] Later, the footprints for the interaction of 16S rRNA with primary r-protein S4 were determined at both temperatures.[66] The S4-specific footprints are different at the two temperatures studied; a conformational rearrangement of 16S rRNA is taking place at the heating of the S4/16S rRNA complex.[66] More recently, the binding constant of S4 to the 5’ domain of the 16S rRNA was determined[54] at 0°C and 37°C in the same study and a four fold difference was observed. The aforementioned studies emphasize how a combination of approaches can result in a better understanding of RNA–protein interactions, and also illustrate the power of chemical footprinting, which was the only biochemical technique that showed the existence of the conformational change of the RNP.

The conformational change from RI to RI* is temperature dependent, the same as the conformational change in the minimal complex S4/16S rRNA. There are other primary r-proteins present in these assembly intermediates, along with some secondary binding proteins. Is their interaction with 16S rRNA also temperature dependent? How is the

www.manaraa.com

conformational change from RI to RI* related to the conformational changes in the minimal complexes? Some of these questions will be addressed in this thesis.

Directed hydroxyl radical probing in the study of the 30S subunit

In the past, directed hydroxyl radical probing has been used mainly to study the nucleic acid environment of proteins in static mature complexes.[27,59,67-73] In the ribosome and ribosomal subunits it helped in the characterization of the RNA surroundings of components such as r-proteins or ligands like tRNA. Both proteins and RNA can be derivatized with Fe(II)BABE (where BABE is (1-(p-bromoacetamidobenzyl) ethylenediaminetetraacetate)), the proteins through cysteine[59] (see Figure 4a and b) and RNA through phosphorothioate.[67] By using the r-proteins as probes, for example the location of r-protein S20 was analyzed in 30S subunit, and helped clarify the controversy on its location.[27] The 16S rRNA elements surrounding S5 in 70S ribosome were also mapped using this technique.[71] Fe(II)BABE was tethered via 5'-phosphorothioate to *in vitro* transcripts, tRNA and tRNA analogs.[67,74]

Recently, directed hydroxyl probing from a derivatized r-protein was used to address conformational changes in the 16S rRNA during 30S subunit assembly.[75,76] The recombinant system for *in vitro* reconstitution makes possible the construction of minimal RNPs that represent different stages of assembly.[36,37] In these RNPs, Fe(II)-tethered S15 protein was incorporated and the changes in the cleavage patterns were used to assess the changes in the rRNA structure. The starting point was the binary complex of 16S rRNA/Fe(II)-S15 which represents one of the initial stages of 30S subunit assembly. More complex RNPs were also probed and the difference in the cleavage patterns was monitored. The differences observed shed light on the

www.manaraa.com

rearrangement of rRNA elements in the presence of a certain r-protein.[75,76] This approach gave insight into the assembly pathway and roles played by r-proteins in this process and showed that directed hydroxyl radical probing can be employed to study rearrangements that occur in nucleic acid-protein complexes during assembly, ligand association or other cellular processes that can be monitored *in vitro*.

Chemical reagents in the study of RNA-protein interactions. Footprinting and directed hydroxyl radical probing

Footprinting and primer extension are techniques used in the studies presented in this dissertation. Footprinting involves chemical modification of nucleotides and allows identification of nucleotides affected by the binding of a protein to RNA.[57] Chemical modification can be selective when base specific probes are employed. Some of the base specific probes interact only with unpaired nucleotides, revealing changes in the secondary structure of the RNA.[57] Hydroxyl radicals cleave the sugar phosphate backbone nonselectively. Diminished reactivity of a nucleotide as a result of protein binding can be due to direct protein-RNA interaction or a conformational rearrangement of RNA.[57] Hydroxyl radicals are not very sensitive to secondary structure and usually reduced reactivity after binding implies protein-RNA contact.[77] Enhanced reactivity after binding implies a ligand induced conformational change. In our studies we are using dimethyl sulfate (DMS) which methylates selectively adenines (at position N1) and cytosines (at position N3), and 3-ethoxy-2-ketobutanal (kethoxal) which modifies reversibly guanines (by forming a cyclic adduct)[57]. Both DMS and kethoxal react only with unpaired nucleotides and are very sensitive to the secondary structure of RNA.[57] By

www.manaraa.com

using the two types of footprinting probes a distinction can be made between direct RNA-protein contacts and conformational changes that take place at some distance from the binding site.

A different footprinting approach is hydroxyl directed probing, which involves generation of hydroxyl radicals by Fenton chemistry only in the area that surrounds a tethered iron ion.[59] This method is very useful in mapping the RNA environment in the vicinity of the probe, which is in our case a protein. Hydroxyl directed cleavage involves a few steps. An r-protein is derivatized at a single position, generally a cysteine, with Fe(II) via the linker 1-(p-bromoacetamidobenzyl)-EDTA (BABE)[59] (Figure 4a). Control experiments ensure that the derivatized protein can still interact with the rRNA in the same way as the wild type protein. The desired RNPs are assembled and the cleavage reaction is performed in the presence of hydrogen peroxide and ascorbic acid. Directed hydroxyl cleavage generates information about the direct surroundings of the protein, and it can be used to systematically map the RNA environment in the vicinity of the protein by tethering the Fe(II) to different positions in the protein.[59] The aforementioned facts eliminate one of the disadvantages of base specific probes, the ambiguity in attributing protections that can appear due to direct contact or conformational changes. The intensity of the cleavage can be used also to estimate the distance between the Fe(II) and the RNA backbone and the large number of cleavage site specific for these probes provide a large number of data points.

Primer extension. The method that makes possible the identification of changes in reactivity of the nucleotides toward different probes or the cleavage sites is primer extension[57, 58] (Figure 5a). A complementary DNA of the RNA molecule of interest is

www.manaraa.com

generated, through a reaction catalyzed by the enzyme reverse transcriptase. Initially, a short DNA oligonucleotide primer is annealed to a certain sequence of the target RNA. The reverse transcriptase is able to extend the DNA complementary to the studied RNA until a modified nucleotide is encountered, when the transcription stops or pauses, and a truncated DNA is generated. To be able to visualize each modified nucleotides, the probing conditions are such that only a fraction of the RNA is modified. Otherwise, only the shortest possible DNA fragment will be generated, since the reverse transcriptase will stop at the first modification. For detection, a radiolabeled nucleotide is incorporated in the complementary DNA. The samples are run on a denaturing sequencing gel, along with the sequencing lanes (A and G), which help in the localization of the modified nucleotides on the RNA. A control lane, which has unmodified RNA is also loaded (K), and natural stops which appear in all lanes are called K-bands (Figure 5b and c). The reactivity of almost every nucleotide can be monitored by this technique. For base-specific probing enhancements and protections are observed in the gel. When the decrease in reactivity of a specific nucleotide is observed, compared to the reactivity of 16S rRNA, a protection takes place and the increase in reactivity indicates an enhancement[57,58] (see Figure 5b). For the hydroxyl directed cleavage experiments, all the bands that are not observed in the control lane or the ones that have a higher intensity identify positions of directed cleavage[59] (Figure 5c).

References

1. Stark, B. C., *Mol. Biol. Rep.* **1995,** 22, (2-3), 95-7.

2. Newman, A., *Curr. Biol.* **1994,** 4, (5), 462-4.

3. Steitz, T. A.; Moore, P. B., *Trends Biochem. Sci.* **2003,** 28, (8), 411-8.

www.manaraa.com

4. Moore, P. B.; Steitz, T. A., *Nature* **2002,** 418, (6894), 229-35.

5. Brodersen, D. E.; Nissen, P., *FEBS J.* **2005,** 272, (9), 2098-108.

6. Williamson, J. R., *Nat. Struct. Biol.* **2000,** 7, (10), 834-7.

7. Leulliot, N.; Varani, G., *Biochemistry* **2001,** 40, (27), 7947-56.

8. Schachman, H. K.; Pardee, A. B.; Stanier, R. Y., *Arch. Biochem. Biophys.* **1952,** 38, 245-60.

9. Blaha, G., Structure of the Ribosome. In *Protein Synthesis and Ribosome Structure*, Nierhaus, K. H.; Wilson, D. N., Eds. Wiley-VCH: Weinnheim, 2004; pp 53-84.

10. Ramakrishnan, V., *Cell* **2002,** 108, (4), 557-572.

11. Noller, H. F.; Baucom, A., *Biochem. Soc. Trans.* **2002,** 30, (Pt 6), 1159-61.

12. Yusupov, M. M.; Yusupova, G. Z.; Baucom, A.; Lieberman, K.; Earnest, T. N.; Cate, J. H.; Noller, H. F., *Science* **2001,** 292, (5518), 883-96.

13. Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H., *Science* **2005,** 310, (5749), 827-34.

14. Brosius, J.; Palmer, M. L.; Kennedy, P. J.; Noller, H. F., *Proc. Natl. Acad. Sci. U.S.A.* **1978,** 75, (10), 4801-5.

15. Noller, H. F.; Woese, C. R., *Science* **1981,** 212, (4493), 403-11.

16. Ludwig, W.; Schleifer, K. H., *FEMS Microbiol. Rev.* **1994,** 15, (2-3), 155-73.

17. Agalarov, S. C.; Selivanova, O. M.; Zheleznyakova, E. N.; Zheleznaya, L. A.; Matvienko, N. I.; Spirin, A. S., *Eur. J. Biochem.* **1999,** 266, (2), 533-537.

18. Samaha, R. R.; O'Brien, B.; O'Brien, T. W.; Noller, H. F., *Proc. Natl. Acad. Sci. U.S.A.* **1994,** 91, (17), 7884-8.

www.manaraa.com

19. Brodersen, D. E.; Clemons, W. M., Jr.; Carter, A. P.; Wimberly, B. T.; Ramakrishnan, V., *J. Mol. Biol.* **2002,** 316, (3), 725-68.

20. Wilson, D. N.; Nierhaus, K. H., *Crit. Rev. Biochem. Mol. Biol.* **2005,** 40, (5), 243-67.

21. Traub, P.; Nomura, M., *Cold Spring Harb. Symp. Quant. Biol.* **1969,** 34, 63-7.

22. Stern, S.; Changchien, L. M.; Craven, G. R.; Noller, H. F., *J. Mol. Biol.* **1988,** 200, (2), 291-9.

23. Stern, S.; Wilson, R. C.; Noller, H. F., *J. Mol. Biol.* **1986,** 192, (1), 101-10.

24. Svensson, P.; Changchien, L. M.; Craven, G. R.; Noller, H. F., *J. Mol. Biol.* **1988,** 200, (2), 301-8.

25. Nowotny, V.; Nierhaus, K. H., *Biochemistry* **1988,** 27, (18), 7051-5.

26. Capel, M. S.; Engelman, D. M.; Freeborn, B. R.; Kjeldgaard, M.; Langer, J. A.; Ramakrishnan, V.; Schindler, D. G.; Schneider, D. K.; Schoenborn, B. P.; et al., *Science* **1987,** 238, (4832), 1403-6.

27. Culver, G. M.; Noller, H. F., *RNA* **1998,** 4, (12), 1471-80.

28. Powers, T.; Daubresse, G.; Noller, H. F., *J. Mol. Biol.* **1993,** 232, (2), 362-74.

29. Golden, B. L.; Hoffman, D. W.; Ramakrishnan, V.; White, S. W., *Biochemistry* **1993,** 32, (47), 12812-20.

30. Davies, C.; Gerstner, R. B.; Draper, D. E.; Ramakrishnan, V.; White, S. W., *EMBO J.* **1998,** 17, (16), 4545-4558.

31. Sayers, E. W.; Gerstner, R. B.; Draper, D. E.; Torchia, D. A., *Biochemistry* **2000,** 39, (44), 13602-13613.

www.manaraa.com

32. Clemons, W. M., Jr.; Davies, C.; White, S. W.; Ramakrishnan, V., *Structure* **1998,** 6, (4), 429-438.

33. Agalarov, S. C.; Sridhar Prasad, G.; Funke, P. M.; Stout, C. D.; Williamson, J. R., *Science* **2000,** 288, (5463), 107-13.

34. Powers, T.; Changchien, L. M.; Craven, G. R.; Noller, H. F., *J. Mol. Biol.* **1988,** 200, (2), 309-19.

35. Held, W. A.; Mizushima, S.; Nomura, M., *J. Biol. Chem.* **1973,** 248, (16), 5720-30.

36. Culver, G. M.; Noller, H. F., *Methods Enzymol.* **2000,** 318, (RNA-Ligand Interactions, Pt. B), 446-460.

37. Culver, G. M.; Noller, H. F., *RNA* **1999,** 5, (6), 832-43.

38. Mizushima, S.; Nomura, M., *Nature* **1970,** 226, (5252), 1214-18.

39. Held, W. A.; Ballou, B.; Mizushima, S.; Nomura, M., *J. Biol. Chem.* **1974,** 249, (10), 3103-11.

40. Nashimoto, H.; Held, W.; Kaltschmidt, E.; Nomura, M., *J. Mol. Biol.* **1971,** 62, (1), 121-38.

41. Nomura, M., *Fed. Proc.* **1972,** 31, (1), 18-20.

42. Held, W. A.; Nomura, M., *Biochemistry* **1973,** 12, (17), 3273-81.

43. Traub, P.; Nomura, M., *J. Mol. Biol.* **1968,** 34, (3), 575-93.

44. Guthrie, C.; Nashimoto, H.; Nomura, M., *Proc. Natl. Acad. Sci. U.S.A.* **1969,** 63, (2), 384-91.

45. Atmadja, J.; Stiege, W.; Zobawa, M.; Greuer, B.; Osswald, M.; Brimacombe, R., *Nucleic Acids Res.* **1986,** 14, (2), 659-73.

www.manaraa.com

46. Atmadja, J.; Brimacombe, R.; Blocker, H.; Frank, R., *Nucleic Acids Res.* **1985,** 13, (19), 6919-36.

47. Brimacombe, R.; Stiege, W.; Kyriatsoulis, A.; Maly, P., *Methods Enzymol.* **1988,** 164, 287-309.

48. Chiaruttini, C.; Milet, M.; Hayes, D. H.; Expert-Bezancon, A., *Biochimie* **1989,** 71, (7), 839-52.

49. Chiaruttini, C.; Expert-Bezancon, A.; Hayes, D.; Ehresmann, B., *Nucleic Acids Res.* **1982,** 10, (23), 7657-76.

50. Expert-Bezancon, A.; Chiaruttini, C., *Methods Enzymol.* **1988,** 164, 310-8.

51. Mougel, M.; Eyermann, F.; Westhof, E.; Romby, P.; Expert-Bezancon, A.; Ebel, J. P.; Ehresmann, B.; Ehresmann, C., *J. Mol. Biol.* **1987,** 198, (1), 91-107.

52. Schwarzbauer, J.; Craven, G. R., *Nucleic Acids Res.* **1981,** 9, (9), 2223-37.

53. Vartikar, J. V.; Draper, D. E., *J. Mol. Biol.* **1989,** 209, (2), 221-34.

54. Gerstner, R. B.; Pak, Y.; Draper, D. E., *Biochemistry* **2001,** 40, (24), 7165-73.

55. Draper, D. E.; Deckman, I. C.; Vartikar, J. V., *Methods Enzymol.* **1988,** 164, 203-20.

56. Mougel, M.; Ehresmann, B.; Ehresmann, C., *Biochemistry* **1986,** 25, (10), 2756-65.

57. Stern, S.; Moazed, D.; Noller, H. F., *Methods Enzymol.* **1988,** 164, 481-9.

58. Moazed, D.; Stern, S.; Noller, H. F., *J. Mol. Biol.* **1986,** 187, (3), 399-416.

59. Culver, G. M.; Noller, H. F., *Methods Enzymol.* **2000,** 318, 461-75.

www.manaraa.com

60. Moore, P. B., The RNA folding problem. In *The RNA World*, Gesteland, R. F.; Cech, T. R.; Atkins, J. F., Eds. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1999; pp 381-401.

61. Holmes, K. L.; Culver, G. M., *J. Mol. Biol.* **2005,** 354, (2), 340-357.

62. Holmes, K. L.; Culver, G. M., *Nat. Struct. Mol. Biol.* **2004,** 11, (2), 179-186.

63. Talkington, M. W.; Siuzdak, G.; Williamson, J. R., *Nature* **2005,** 438, (7068), 628-32.

64. Brewer, L. A.; Noller, H. F., *Biochemistry* **1983,** 22, (18), 4310-5.

65. Sapag, A.; Vartikar, J. V.; Draper, D. E., *Biochim. Biophys. Acta* **1990,** 1050, (1-3), 34-7.

66. Powers, T.; Noller, H. F., *J. Biol. Chem.* **1995,** 270, (3), 1238-42.

67. Joseph, S.; Noller, H. F., *Methods Enzymol.* **2000,** 318, 175-90.

68. Samaha, R. R.; Joseph, S.; O'Brien, B.; O'Brien, T. W.; Noller, H. F., *Proc. Natl. Acad. Sci. U.S.A.* **1999,** 96, (2), 366-370.

69. Newcomb, L. F.; Noller, H. F., *Biochemistry* **1999,** 38, (3), 945-51.

70. Newcomb, L. F.; Noller, H. F., *RNA* **1999,** 5, (7), 849-55.

71. Culver, G. M.; Heilek, G. M.; Noller, H. F., *J. Mol. Biol.* **1999,** 286, (2), 355-64.

72. Heilek, G. M.; Noller, H. F., *Science* **1996,** 272, (5268), 1659-62.

73. Heilek, G. M.; Noller, H. F., *RNA* **1996,** 2, (6), 597-602.

74. Joseph, S.; Whirl, M. L.; Kondo, D.; Noller, H. F.; Altman, R. B., *RNA* **2000,** 6, (2), 220-32.

75. Jagannathan, I.; Culver, G. M., *J. Mol. Biol.* **2004,** 335, (5), 1173-1185.

76. Jagannathan, I.; Culver, G. M., *J. Mol. Biol.* **2003,** 330, (2), 373-383.

www.manaraa.com

77. Powers, T.; Noller, H. F., *RNA* **1995,** 1, (2), 194-209.

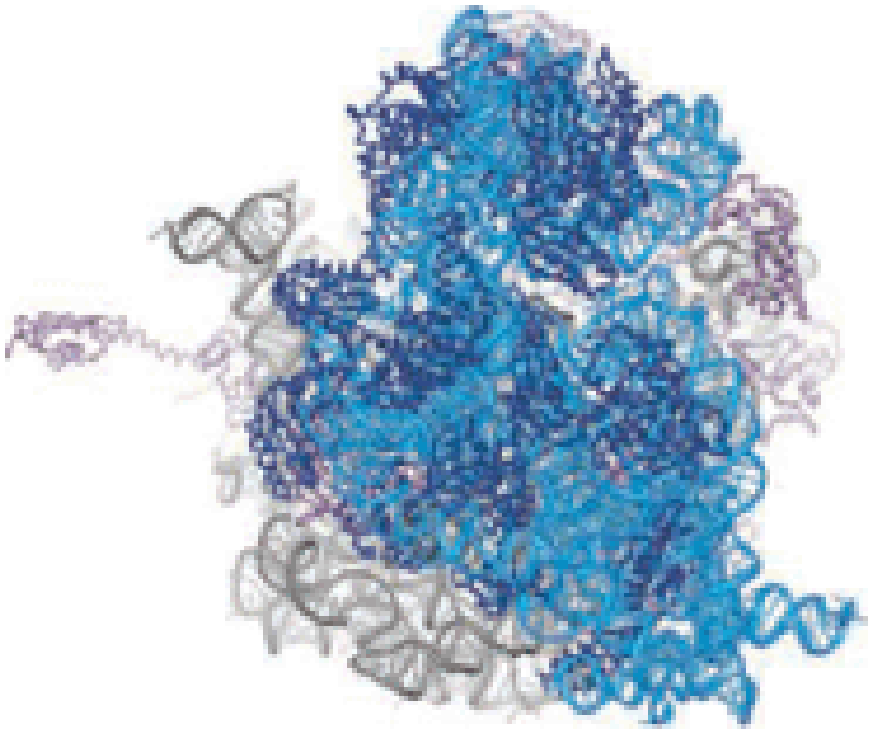
78. Cannone, J. J.; Subramanian, S.; Schnare, M. N.; Collett, J. R.; D'Souza, L. M.; Du, Y.; Feng, B.; Lin, N.; Madabusi, L. V.; Muller, K. M.; Pande, N.; Shang, Z.; Yu, N.; Gutell, R. R., *BMC Bioinf.* **2002,** 3, 2.

79. Grondek, J. F.; Culver, G. M., *RNA* **2004,** 10, (12), 1861-1866.

www.manaraa.com

Figure 1. Crystal structure of the Escherichia coli ribosome (70 S). The 30S (small) subunit is shown in blue for the 16S rRNA and dark blue for the ribosomal proteins, in front. The 50S (large) subunit is shown in gray for the 23S rRNA and the proteins in purple, in the back.

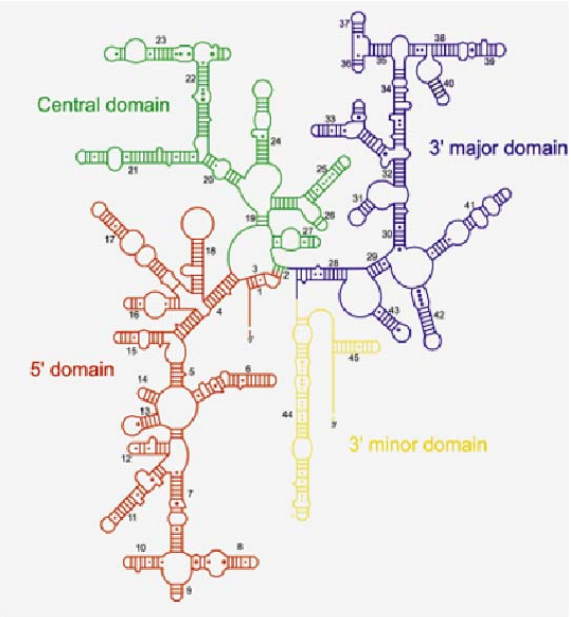
www.manaraa.com



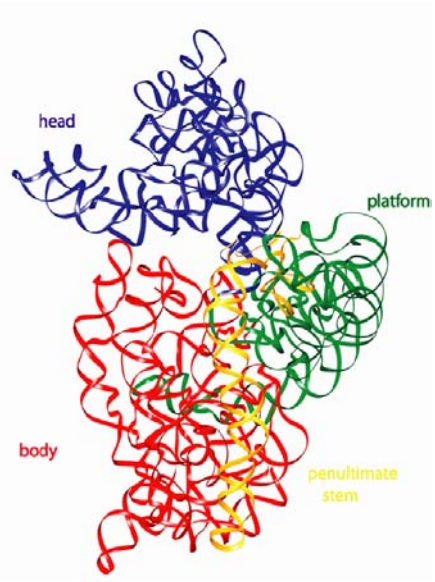
www.manaraa.com

Figure 2.16S rRNA and the 30S subunit organization. (a) Secondary structure of 16S rRNA with its domains in different color. Red the 5' domain, green the central domain, blue the 3' major domain and in yellow the 3' minor domain.[78] (b) Tertiary structure of the 16S rRNA with its different three dimensional parts in the color corresponding to the domain in the secondary structure. Head is in blue, platform in red, body in green and penultimate stem in light grey.[13] (c) *In vitro* assembly map of 30S subunit with proteins binding to the different domains in the respective colors.[38,79] The proteins in the dark gray box are primary binding proteins, the ones in the light gray box are secondary binding proteins and the proteins in the white area are tertiary binding proteins.

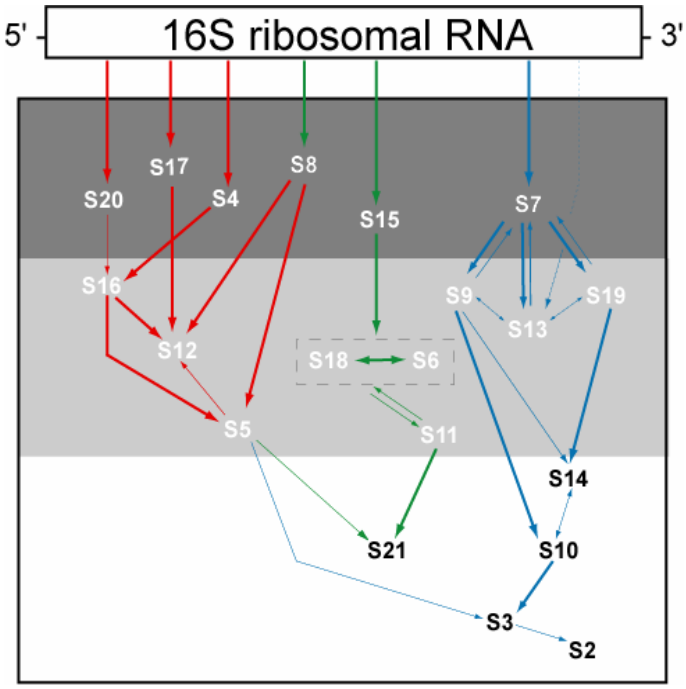
www.manaraa.com



(a)



(b)



(c)

www.manaraa.com

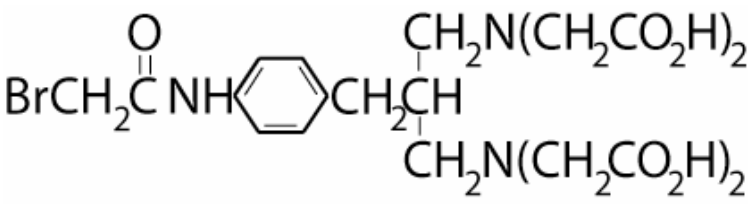
Figure 3. Three dimensional structure of 16S rRNA from 30S subunit with the primary r-proteins.[13] 16S rRNA is showed in gray, S4 in green, S7 in red, S8 in pink, S15 in lime yellow, S17 in dark purple and S20 in light blue.

www.manaraa.com

www.manaraa.com

Figure 4. Directed hydroxyl radical probing. (a) Structure of 1-(p-bromoacetamidobenzyl)-ethylenediaminetetraacetic acid (BABE) the linker through which Fe(II) is tethered. (b) Scheme for directed hydroxyl radical probing of RNPs.

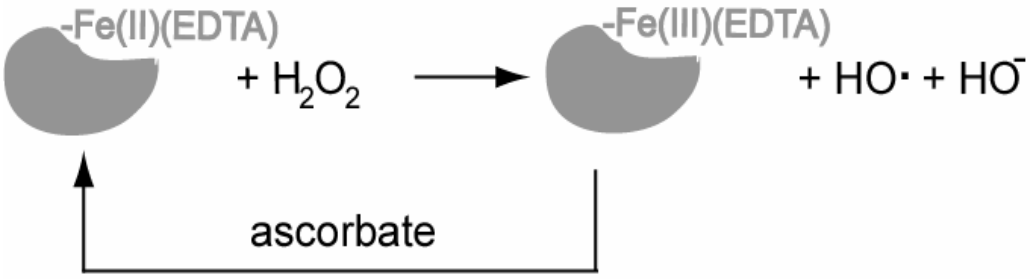
www.manaraa.com



(a)

1. Generate cysteine containing proteins

2. Assess the availability of the cysteine for derivatization
3. Attach Fe(II)EDTA via the cysteine
4. Determine that the modified protein retains function
5. Assemble and purify ribonucleoprotein particles
6. Initiate hydroxyl radical probing



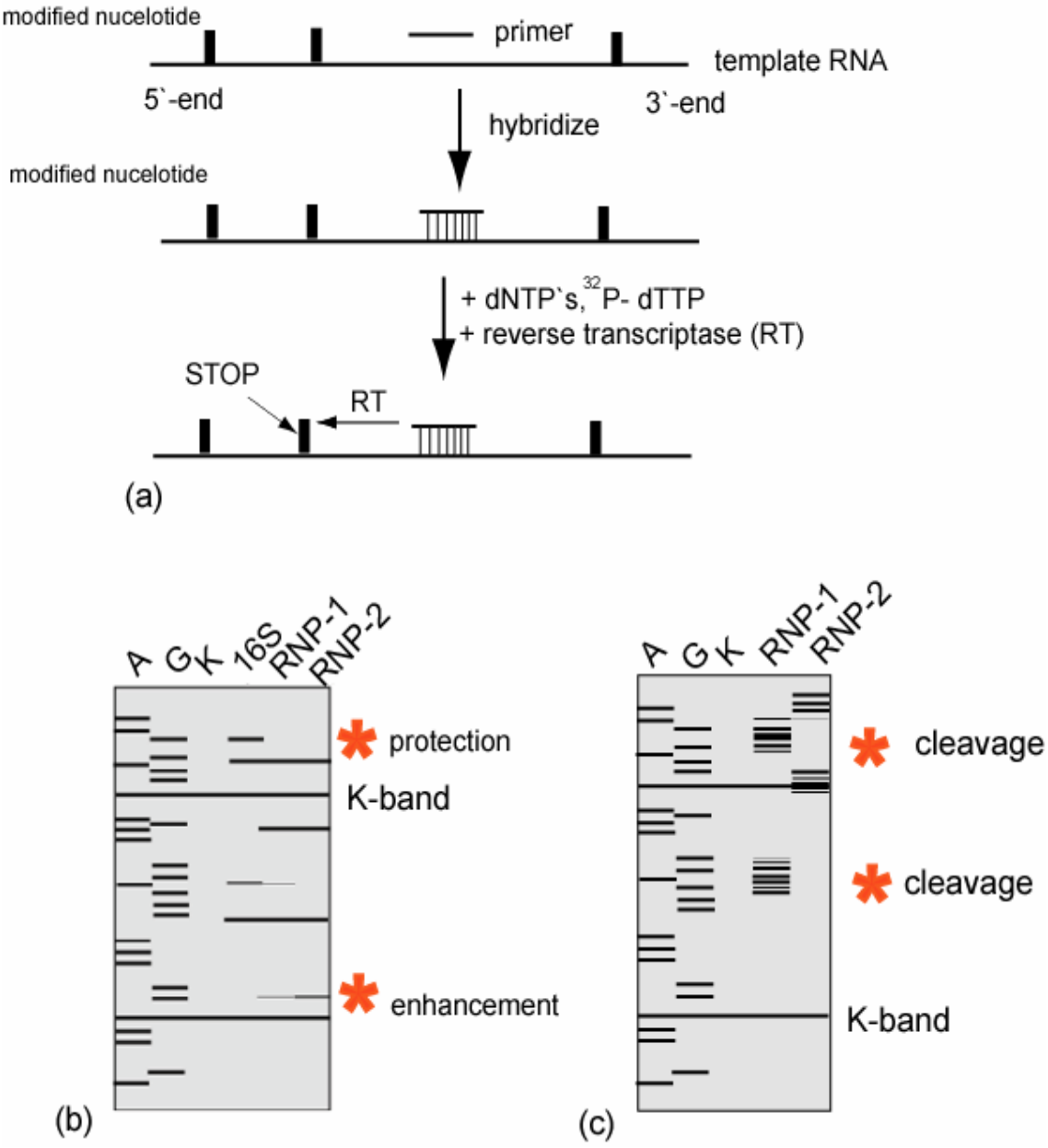
7. Identify the sites of cleavage by primer extension

(b)

www.manaraa.com

Figure 5. Primer extension. (a) Scheme for primer extension. (b) Schematic gel for base-specific footprinting. Enhancements, protections and K-bands are indicated. (c) Schematic gel for directed hydroxyl radical probing. Cleavage sites and K-bands are indicated.

www.manaraa.com



www.manaraa.com

Chapter VI. Temperature-dependent RNP conformational rearrangements: analysis of binary complexes of primary binding proteins with 16S rRNA

Laura-M. Dutcă, Indu Jagannathan, Joel F. Grondek, Gloria M. Culver

Manuscript in preparation

Abstract

Ribonucleoprotein particles (RNPs) are important components of all living systems, and the assembly of these complexes is an intricate process. The 30S ribosomal subunit is composed of one large RNA (16S rRNA) and 21 ribosomal proteins (r-proteins). 30S subunit assembly has been shown to involve sequential binding of r-proteins and conformational changes of 16S rRNA. *In vitro* studies have revealed that assembly of 30S subunit is a highly temperature dependent process. Given these observations, a systematic study of the temperature dependence of 16S rRNA architecture in individual complexes with five primary proteins (S7, S8, S15, S17 and S20) was performed. Our data suggest that some temperature-dependent conformational changes occur and are consistent with downstream assembly events. As expected, all r-proteins can bind 16S rRNA at low temperature. However not all r-proteins/16S rRNA complexes undergo temperature-dependent conformational rearrangements. Some RNPs acquire the same conformation regardless of temperature, others show minor adjustments in 16S rRNA conformation upon heating, and finally others undergo significant temperature-dependent conformational changes. Some of the architectures achieved in these temperature-dependent conformational rearrangements are likely required for further

www.manaraa.com

assembly of the secondary and tertiary binding r-proteins. The differential interaction of 16S rRNA with r-proteins illustrates a means for controlling the sequential assembly pathway for complex RNPs and may offer insights into aspects of RNP assembly.

Introduction

The interaction between RNA and proteins to form functional ribonucleoprotein particles (RNPs) is a very exacting process. Often, binding of protein to RNA involves major conformational changes in one or both molecules. Dynamic or disordered elements of the binding partners can adopt a defined conformation in the complex.[1,2] This process is called induced fit, and it is believed to be a major component in the assembly of multicomponent RNPs. These changes can be dramatic, from completely disordered in the free form to strictly constrained in the complex, or somewhat more subtle, such as changes in domain orientation.[1,2] The presence of multiple domains in one of the partners also makes possible sequential binding to these different domains in a binding cascade. Moreover, conformational rearrangements can be propagated throughout the molecule and are not limited to the binding interface. While the understanding of RNA-protein interactions has been greatly enhanced by advances in RNP crystallography, a detailed view of conformational changes during RNP assembly is still lacking. Systematic studies using a well characterized model system, the 30S ribosomal subunit will advance our understanding of events central to RNP assembly. The *Escherichia coli* (*E. coli*) 30S ribosomal subunit is composed from 16S ribosomal RNA (rRNA) and 21 ribosomal proteins (r-proteins), and it can be a rich source of information for the student of RNA-protein interactions. The crystal structure of the ribosome from *E. coli* was determined recently,[3] and detailed structures of the individual ribosomal subunits from multiple

www.manaraa.com

organisms are also available.[4-6] These structures are very useful in analyzing assembly events, as they represent an end point for the assembly process. Additionally, the structures of some of the free r-proteins have also been determined.[7-12] These findings allowed comparisons of free and RNA bound r-proteins, leading to inferences about changes in r-protein structure, as a result of RNP assembly. However, very few detailed structures of naked 16S rRNA are available and thus similar inferences about RNA conformational changes during ribosome assembly are lagging.

Some advances in understanding RNA conformational changes have arisen from studies of *E. coli* 30S subunit. One reason for these advances is the ability to reconstitute the 30S subunit in a functional conformation *in vitro* from its isolated components.[13] RNPs of different complexities can be readily formed and this system has allowed the elucidation of certain aspects of the formation of 30S subunit architecture. Distinct three-dimensional structures arise when the four secondary structural domains of 16S rRNA are complexed with the appropriate r-proteins. In the 30S subunit, the 5' domain of 16S rRNA forms the body of the 30S subunit, the central domain folds into the platform, the 3' major domain forms the head and the 3' minor domain is mainly folded into the penultimate stem.[3,14] Thus the secondary structure of 16S rRNA is influential in determining 30S subunit architecture.

The sequential binding of the r-proteins to 16S rRNA is a critical step in orchestrating formation of functional 30S subunits. Traditionally, the r-proteins have been categorized into three assembly classes, as indicated in the *in vitro* assembly map (Figure 1a).[15,16] The r-proteins that bind directly and independently to 16S rRNA are classified as primary, and they are S4, S7, S8, S15, S17 and S20. The secondary binding

www.manaraa.com

proteins, S5, S6, S9, S11, S12, S13, S16, S18 and S19 bind 16S rRNA after the assembly of at least one primary protein, while the tertiary binding proteins, S2, S3, S10, S14 and S21, require association of at least one primary and one secondary r-protein for their binding. The r-proteins can also be slightly differently classified based on the dynamics of their association with 16S rRNA during assembly at different temperatures.[17] The r-proteins were classified in the following assembly kinetic classes: early binders (S4, S6, S11, S15, S16, S17, S18, S20), mid binders (S7, S8, S9, S13, S19), mid-late (S5 and S12) and late (S2, S3, S10, S14, S21) binders by following the emergence of their footprints during assembly[17].

During *in vitro* reconstitution of 30S subunits the primary and secondary binding r-proteins associate early in assembly. A 16S rRNA containing RNP, known as RI (Reconstitution Intermediate), containing these r-proteins is formed at low temperature. However, a temperature dependent step is essential for the formation of a second intermediate, RI*, with virtually the same composition as RI. A large compaction of the RNP occurs during this activation event, suggesting that the 16S rRNA and these r-proteins undergo a conformational rearrangement in response to the temperature activation, and that this change is required for assembly to proceed to completion.[18-21] In addition, *in vivo* 30S subunit assembly defects are associated with cold-sensitivity and assembly intermediates similar to those observed *in vitro* have been isolated.[22-24] Thus it appears that analysis of temperature-dependent conformational changes *in vitro* may have some bearing on the *in vivo* assembly pathway.

Analysis of 30S subunit assembly in the presence of all or many of the r-proteins, have revealed global trends, without dissecting the role of each of the individual r-

www.manaraa.com

proteins. The presence of all r-proteins allows the concerted changes in the conformation of 16S rRNA, but by necessity obscures the contribution of individual proteins. For example, in the study of the temperature dependent dynamics of 30S assembly, the r-proteins were classified in different kinetic classes based on their footprints observed in previous studies of less complex RNPs.[17] For a few primary binding proteins only some of their footprints were observed. For example, there are approximately 15 protections and enhancements specific for S15 in a minimal RNP,[25] only five of them could be assigned during the ensemble assembly experiment.[17] Similarly, in the minimal complex S7/16S rRNA S7 footprints more than 60 nucleotides,[26] but only about half of this number were attributed to S7 in the ensemble studies.[17] Thus, while these bulk approaches can be illuminating, many changes can be masked or invisible, and further analysis of smaller RNPs may be necessary to fully dissect the changes during assembly of complex systems.

Conformational changes play an important role in the assembly of the 30S ribosomal subunit. The binding of r-proteins involves conformational changes of 16S rRNA at the RNA–protein interface, but they can also bring about conformational changes at some distance from this interface. As mentioned, the transition from RI to RI* involves a large conformational change and this can be facilitated by increased temperature and some analysis of these changes has been performed in a complete assembly reaction.[20,21] It could be of great interest to determine more exactly which r-proteins contribute to these specific conformational changes during the assembly process, to further our understanding of the roles of r-proteins in orchestrating the architectural changes. This approach has proven useful in analyzing the interaction of 16S rRNA with

www.manaraa.com

an r-protein, S4, as function of temperature.[27] S4 is a primary multidomain binding protein[9] (Figure 1a) that is considered an assembly initiator,[28] and it is a component of RI and RI*.[18] A temperature-dependent conformational rearrangement of 16S rRNA in the presence of S4 was observed when the complexes were formed at different temperatures (0, 30 and 42°C).[27] These changes were revealed by differences in the chemical modification pattern of 16S rRNA in these complexes. This approach may be particularly fruitful now as the more recently available structures of 30S subunits may make it possible to better understand the implications of temperature influence on the r-protein-rRNA interaction, and subsequently on assembly.

The structures of 30S subunits have revealed that some of the r-proteins bind multiple domains of the 16S rRNA and some have multiple domains themselves,[29] raising the possibility that the *in vitro* temperature requirement can be used to deconvolute the interaction of these multidomain partners. While the secondary structure of rRNA is usually well defined even in the free form, tertiary structures specific for functional 30S subunit are achieved only after r-proteins bind. To our knowledge, besides the paper of Powers et al. (1995) there are few studies in which protein-RNA interactions at different temperatures are dissected by structural methods. Some studies have shown a dependence of the kinetics of RNA-protein interactions on temperature. Additionally, there are a few studies in which small differences in the RNA-protein binding constants at different temperatures were observed,[30] but no detailed structural analysis of the complexes was undertaken. Moreover, such studies of large RNA-protein complexes are particularly lacking and one could imagine that temperature is an important factor for long range tertiary interactions adopted by large RNAs. Another possible role for the

www.manaraa.com

observed temperature effects on RNA-protein interactions could be control of sequential binding of proteins in complex RNPs, like the ribosomal subunits. Therefore it might not be surprising that *in vitro* assembly of 30S subunits is temperature dependent and that the interaction of at least one r-protein, S4, with 16S rRNA is influenced by temperature. The data presented in this manuscript reveals that some of the possible mechanisms appear to be involved in r-protein/16S rRNA interactions. However, there appear to be distinct differences in how these RNPs form and the influence of temperature on their conformation.

Materials and methods

16S rRNA/r-protein complex formation. The complexes were prepared from 16S rRNA and the 30S subunit recombinant r-proteins isolated as described previously.[31-33] The buffers used at the formation of the RNPs are: reconstitution A minus buffer (RA-) which is 20 mM K^+-Hepes (pH 7.6), 20 mM $MgCl_2$, 6 mM β-mercaptoethanol; Reconstitution A plus buffer (RA+) which has the same composition as RA-, plus 330 mM KCl. The complexes were formed as follows: natural 16S rRNA in RA- was incubated at 42°C for 15 minutes, followed by 10 minutes on ice prior to complex formation. 40 pmoles of 16S rRNA were mixed with 200 or 240 pmoles of the appropriate r-protein, and the final KCl concentration was adjusted to 330 mM, by using the appropriate ratios of RA+ and RA-, and taking into account that the protein solutions are 1 M in KCl. The reaction mixture was incubated at the desired temperature, 0°C or 42°C for 1 hour, or for the shifted complex, 30 min at 0°C and 30 minutes at 42°C. Two

www.manaraa.com

samples containing only 16S rRNA were also incubated at 0°C or 42°C, for comparison. All samples were then incubated on ice for ten minutes, before probing.

Chemical probing and primer extension analysis of the 16S rRNA/r-protein complexes. Chemical probing of 16S rRNA, and the RNPs with kethoxal and DMS was performed as previously described,[31,34] on ice. The probing times were: kethoxal, 60 min and dimethyl sulfate (DMS), 120 min. Primer extension was performed essentially as described.[31,34]

Results

In an attempt to dissect the influence of temperature on RNA-protein interactions, a footprinting study of the complexes of 16S rRNA and individual primary r-proteins formed at different temperatures was undertaken. Complexes between individual primary binding r-proteins and 16S rRNA are formed at different temperature and the interactions are analyzed by chemical modification and primer extension. This approach should allow temperature-dependent conformational changes to be revealed. This systematic analysis of the independent interactions of all of the primary binding r-proteins will allow a better understanding of the rRNA/r-protein interactions and of the importance of temperature in the assembly of the 30S subunit.

Complex formation and chemical probing of binary RNPs. Individual complexes of natural 16S rRNA with the recombinant primary binding proteins S7, S8, S15, S17 and S20 were formed by incubating the reaction mixture at either 0^{o}C or 42^{o}C and a third complex was formed by incubating the reaction mixture first at 0^{o}C and then at 42^{o}C, and herein will be referred to as "shifted" complex. Once complexes had been formed, all particles were placed on ice and probing was performed at low temperature.

www.manaraa.com

This approach will allow the detection of r-protein facilitated temperature-dependent differences in 16S rRNA architecture and for the "shifted" complex will reveal if either the low or high temperature interactions are predominant. The reactivity of 16S rRNA in these complexes and of naked 16S rRNA incubated at the appropriate temperature, toward the base specific probes dimethyl sulfate (DMS) and kethoxal was investigated by primer extension analysis. In the previous work on the S4/16S rRNA RNP, no temperature-dependent differences were observed when the RNA backbone was probed[27] and our findings for other RNPs are consistent with this earlier work. The reactivities observed for naked 16S rRNA at 42°C and complexes formed at 42°C were very similar to the ones previously published for similar conditions.[25,26,35] In some cases the reactivities observed at 0°C are clearly distinct from those observed at 42°C. Interestingly, only some of the primary binding r-proteins/16S rRNA RNPs revealed temperature dependent conformational changes. Our data reveal that the footprints observed for the shifted complex were essentially the same as those observed for the complexes formed at 42°C, indicating that the particles formed at low temperature can transition from one conformation to another (but not the opposite). Given the similar probing patterns in the shifted complexes and those formed at 42°C these two sets of RNPs will be discussed as one. The results indicate that the primary binding protein/16S rRNA RNPs can be classified into three distinct groups as regards temperature-dependent conformational rearrangements. One class reveals a large temperature-dependence, as previously reported for S4.[27] A second class reveals slight temperature-dependence, where most of the footprints are observed at low temperature, but the intensity is not fully reached until after heating. Somewhat surprisingly, the proteins from the third class

www.manaraa.com

reveal virtually no temperature-dependence. Thus it appears that not all primary r-protein/16S rRNA complexes undergo temperature-dependent conformational rearrangements. Data for each of the r-protein/16S rRNA complexes will be discussed below. Throughout this manuscript, when we refer to previous footprints these were determined for complexes of r-proteins with 16S rRNA that were formed at high temperature (a comprehensive low temperature analysis is lacking).

S20/16S rRNA complexes. X-ray crystallographic studies of 30S subunits reveal that S20 is one of the few r-proteins that interacts with two different domains of 16S rRNA, the 5' domain (body) and H44 of the 3' minor domain[3, 36](Figure 1b). This is consistent with all of S20 footprints being localized in the 5' domain and in H44[35]. A model of 5' to 3' assembly and the positioning of helix 44 across the body in the small subunit, might suggest differential interaction of S20 with these domains. However, no differences in the footprints were observed between the S20/16S rRNA complexes regardless of the temperature at which they are formed (see Table 1, Figure 2a-d, Figure 3). Thus, it appears that in the minimal binary particle S20 and 16S rRNA interact in a temperature independent manner, and for the S20/16S rRNA complex the "desired" conformation is established even at 0^{o}C. These results are in marked contrast to those previously reported for the S4/16S rRNA RNP.[27]

S17/16S rRNA complexes. Structural studies revealed that in the context of 30S subunits, r-protein S17 makes contacts with helix 11 of the 5' domain and helices 20, 21, 22 of the central domain[3, 36](Figure 1b). However in the minimal S17/16S rRNA particle which has been used to identify S17-dependent footprints, the chemical footprints observed for S17 are present almost exclusively (all but 3) in helix 11.[35] Our results

www.manaraa.com

reveal that the footprints, which are all protections, are the same no matter the temperature of S17/16S rRNA complex formation (Table 1, Figure 2 e, f, Figure 3). Thus, similar to what was observed with S20 (see above), temperature seems to have no effect on the footprints of r-protein S17 on 16S rRNA. These results suggest that r-protein dependent organization of portions of the 5' domain can occur in a single step and that the temperature of complex formation has no obvious effect on this interaction.

S15/16S rRNA complexes. The footprints specific for S15 in the minimal RNP are localized in helices 22 and 23[25] and these are very consistent with the S15 binding site, the three way junction between H20, H21 and H22 revealed in the full 30S subunits (Figure 1b).[3, 36] Thus, in minimal and more elaborate particles,[37] S15 interactions are restricted to the central domain of 16S rRNA. For the S15/16S rRNA complexes the results are different than those observed for S20 and S17; slight temperature-dependent conformational adjustments were observed for the S15/16S RNP (Table 1, Figure 2 g, h, Figure 3), and all of the temperature dependent footprints follow a similar pattern (Figure 3). All of the protections (temperature dependent or not) appear at 0^{o}C, but four of these develop in intensity at higher temperature. Only one temperature dependent footprint is observed in helix 22, where the majority of the crystal contacts between S15 and 16S rRNA are observed[3] (Figure 3). The majority of these changes in reactivity of nucleotides due to altered temperature are present in helix 23. Four temperature dependent footprints are observed at nucleotides within helix 23, which is not involved in direct RNA-protein contacts in the crystal structure of 30S subunits[3]. These results suggest that binding occurs at low temperature but that the association of S15 with 16S rRNA is further accommodated at higher temperature.

www.manaraa.com

On the three dimensional structure of 16S rRNA from the 30S subunit, the protections coming from S15 are oriented toward the head of 30S subunit (Figure 4a). The enhancements brought by the binding of S15 are also located towards the head of the 30S subunit. The enhancements which are not temperature dependent are more toward the exterior of the 30S subunit while the ones which emerge at 42^0C are toward the interior (Figure 4b). Interestingly, it appears that the temperature-dependent and temperature independent footprints are differentially clustered within the 30S subunit (Figure 4a, b).

S8/16S rRNA complexes. The footprints specific for S8 in the minimal particle are present throughout the central domain and a few are observed in the 5' domain.[25] More precisely the 530 loop, 570 region, helices 20, 21 and 23, and also the 820 and 860 regions are footprinted by S8.[25] S8 footprints near domain junctions both for the 5' and central domain, and for the central domain and the 3' domain (Table 1, Figure 2 i-n, Figure 3). No temperature dependent changes are observed in the 3-way helical junction, H20/H21/H22, in the S8/16S rRNA complex (Figure 3). In the RNP containing S8 and 16S rRNA formed at 0^oC, many of the enhancements and protections specific for the binding of S8 at 42^oC are observed (Figure 2 i-n, Figure 3). However, most nucleotides in helix 23 (Figure 2k, l), 530 and 570 loops (Figure 2i, j), which are footprinted by S8, reveal a temperature dependent requirement for attaining the full extent of footprinting (Figure 3). Hence the majority of the S8 specific footprints are not as intense at 0^oC as at higher temperature. The largest temperature dependent differences in the reactivities are observed for the nucleotides from helix 26 and the 860 region (Figure 2m, n, Figure 3c).

www.manaraa.com

These results suggest that a conformational rearrangement of the S8/16S rRNA complex may be involved in organizing the more 3' elements of the S8 binding site.

The S8 footprints which reveal differential temperature-dependent conformational changes are also somewhat clustered in the mature 30S subunit. On the three dimensional structure of 16S rRNA from 30S subunit the footprinting classes almost appear to be layered (Figure 4c and d). All the protections that appear at 0°C are clustered, and they are in the lower part of the 30S subunit. The protections that emerge only at 42°C are also clustered and are localized more toward the head of the 30S subunit (Figure 4c). The protections that appear at 0°C and achieve a higher level of footprinting at 42°C are grouped and are somewhat in between the other two sets. Thus it appears that there is a relative spatial context to the conformational rearrangement associated with the S8/16S rRNA particle. The aforementioned results may reflect a primary binding event followed by an adjustment of 16S rRNA in the S8/16S rRNA complex at higher temperature.

S7/16S rRNA complexes. S7 nucleates the assembly of the head of the 30S subunit, by binding to two multiple-stem junctions of the 3' domain of 16S rRNA, H28/H29/H43 and H29/H30/H41/H42 (Figure 1b).[3, 36] Consistent with its RNA interactions in the 30S subunit, binding of S7 has been shown to affect the reactivity of many 16S rRNA nucleotides in footprinting experiments.[26] Our data indicate that the S7/16S rRNA RNP undergoes extensive temperature-dependent conformational rearrangements (Table 1, Figure 2o-s, Figure 3). Large differences are observed between the reactivities of 16S rRNA nucleotides in complex with S7 when the RNP is formed at either low or high temperature. For all the regions footprinted by S7 temperature-dependent alterations in reactivity are observed, suggesting that conformational

www.manaraa.com

rearrangements are prevalent for this complex (Figure 3c). When the S7/16S rRNA containing RNP is formed at low temperature only approximately 16% of the high temperature footprints are detected. Interestingly, there is a correlation between footprints that are observed at 0°C and direct contacts between S7 and 16S rRNA that are apparent in the 30S subunit.[3] In particular, at 0°C, protections and enhancements were observed in the 1330 region (mostly enhancements) and 1350/1370 stem-loop structure (only protections) (Figure 2q-s, Figure 3a), but most of the other expected footprints are incomplete or absent. (Direct contacts between S7 and the 1350/1370 loop of 16S rRNA are present in the structure of the 30S subunit.) Strong temperature dependence is observed at the three way junction H28/H29/H43 and the multiple-stem junction H29/H30/H41/H42, suggesting that these elements become associated with S7 as a consequence of a conformational rearrangement (Figure 3c).

The majority of the S7 footprints that appear at 0°C are grouped together on the three dimensional structure of 16S rRNA from 30S subunits[3] (Figure 5c and d). They are localized in the region of 16S rRNA that is near the N-terminus of S7. The protections that appear at 42°C are more dispersed; nonetheless many of them are clustered along one region of the head (Figure 5c). There is also a trend that can be related to the proximity of the sites to S7 and the extent of temperature dependence observed: the enhancements which are more proximal to S7 are initially observed at 0°C and become more intense at 42°C, while the ones that are more distal from S7 mostly appear only in the complex formed at 42°C (Figure 5d). Thus it appears that S7/16S rRNA undergoes an extensive temperature-dependent conformational rearrangement and that this rearrangement is consistent with the architecture of the 30S subunit.

www.manaraa.com

Discussion

The results presented in this manuscript and those published earlier by Powers and Noller[27] clearly illustrate that the interaction of primary binding r-proteins with 16S rRNA can be significantly influenced by temperature. However, the results herein demonstrate that not all the primary binding r-protein/16S rRNA particles undergo temperature-dependent conformational rearrangements. Since the previous studies that revealed the changes for the S4/16S rRNA complex were performed prior to the determination of the 30S subunit structure, we will revisit these data to provide a full picture of the temperature-dependent conformational changes associated with complexes of 16S rRNA and primary binding r-proteins. The study of these relatively simple RNPs, in isolation from the remaining small subunit components, has allowed a detailed analysis of their specific interactions. These studies allow insight into multiple mechanisms of primary binding protein interaction with 16S rRNA and underscore the complexity of 30S subunit assembly and RNP formation in general.

The 16S rRNA/r-protein RNPs can be classified in three categories based on the effect of temperature on their conformation: 16S rRNA/r-protein RNPs for which the conformation of 16S rRNA is not influenced by temperature (S17 and S20), 16S rRNA/r-protein RNPs that show some temperature dependence (S8 and S15) and the last type, 16S rRNA/r-protein RNPs whose conformation shows a marked dependence on temperature (S4 and S7) (Figure 3, Figure 6). This suggests that the assembly of 16S rRNA containing RNPs can occur at distinct stages and that some of these RNPs can progress from one conformation to another in a temperature dependent manner, while others appear to be less dynamic.

www.manaraa.com

Overall, there is a good correlation between protein binding sites and temperature-dependent conformational rearrangements of the RNP. The r-proteins S17 and S20 for which RNP conformation is not influenced by temperature, bind to the 5' domain. S8 and S15 which bind in the central domain of 16S rRNA form RNPs that show some temperature dependence. Lastly, S4 and S7, whose RNPs show the highest temperature dependence, and are considered assembly initiators,[28] bind to the 5' domain and 3' major domain, respectively. The presence of temperature dependent stages in the formation of RNPs and the conformational rearrangements of 16S rRNA during the interaction with some of the primary r-proteins (S4 and S7) suggest a temperature-dependent induced fit mechanism. Also, it is still likely that induced fit can occur in the low temperature binding event. An induced fit mechanism was observed for the binding of S15 to 16S rRNA,[2] yet few temperature dependent conformational changes are observed with this RNP. 16S rRNA seems capable to interact differentially with the primary binding r-proteins, and thus differential interaction with the other r-proteins, is likely as well.

For the RNPs that display strong temperature-dependent conformational rearrangement, a concern might be whether binding occurs at low temperatures or if binding constants are the same at various temperatures. While there are only a few studies in which binding of r-proteins to 16S rRNA are analyzed as a function of temperature, those that have been done are supportive of association of r-proteins with 16S rRNA at low temperatures.[38-41] Binding constants were determined at 0^{o}C and 42^{o}C for S4[38] and S8. The binding constants for S4 association with the appropriate sub-fragments of 16S rRNA have the same order of magnitude (10^{7}) but the value is about four times lower at low temperature than observed at high temperature. For r-protein S8

www.manaraa.com

the constants were reported for binding to full length 16S rRNA and the binding constant at low temperature was about three times lower than at higher temperature, but again the same order of magnitude was observed. The binding rates during assembly for all of the 30S subunit r-proteins except S7 (due to technical problems), were determined by pulse chase quantitative mass spectrometry at 15°C and 40°C.[41] While for some primary binding r-proteins temperature-dependent rate differences are observed, nonetheless binding is present at 15°C, and the rates are consistent with the time of complex formation in our experiments. In addition, *in vitro* 30S subunit assembly studies also indicate that primary and secondary binding r-proteins can bind at low temperature, as an intermediate containing 16S rRNA and these r-proteins is readily detectable.[18,19] Based on these findings it is clear that primary binding r-proteins can associate with 16S rRNA at low temperatures. Thus, these temperature-dependent footprinting changes are not likely due only to association events, but most probably reflect RNP conformational changes.

R-proteins that bind to 16S rRNA in a temperature independent stage. Our findings for S17 and S20 are consistent with other studies looking at assembly dynamics in the context of all the small subunit r-proteins.[17,20,21] S17 and S20 are the only r-proteins that have all of the corresponding footprints in one kinetic class, and they are classified as expected, as early binders.[17] These data are also consistent with a model for association of r-proteins with 16S rRNA in a 5' to 3' manner. The main footprints for both of these r-proteins lie in the 5' domain, and would be expected to bind in an early assembly event. Thus it appears that S17 and S20 have similar properties when studied in isolation or in ensemble studies.

www.manaraa.com

R-proteins that undergo temperature dependent conformational rearrangements with 16S rRNA. S15 is a small 16S rRNA binding protein, for which we observed only a few temperature dependent footprints in the S15/16S rRNA complex. Virtually all of the footprints are observed at 0^{o}C however, some change in intensity at higher temperature (Figure 3). Also, many of its footprints coincide with direct RNA-protein contacts in the 30S subunit.[3,25] The structure of free S15 has been determined by both NMR[11] and X-ray crystallography and is very similar to the structure of S15 in the 30S subunit,[3] suggesting that S15 can obtain its structure in the absence of 16S rRNA. Binding of S15 to 16S rRNA has been shown to result in a large conformational rearrangement of the RNA,[2] which occurs independently of temperature, both in a small fragment or the full length 16S rRNA. Hence our findings are consistent with a single step assembly event for the S15/16S rRNA complex.

The r-protein S15 is one of the proteins were the advantages of studying minimal complexes are obvious. Only five footprints specific for S15 were identifiable in the assembly dynamics study,[17] while in our experiments we were able to assess majority of the S15 dependent footprints. Additionally, while we observe many footprints at 0^{o}C, only one footprint was attributable to S15 at that temperature, when all the r-proteins are present.[17] Thus our experiments give a more detailed understanding of the S15 binding process.

The primary r-protein S8 interacts with an extensive region of the central domain of 16S rRNA and it is a mid-binder from kinetic point of view.[17] In this study it was shown that the S8/16S rRNA complex undergoes a temperature-dependent conformational rearrangement. The temperature dependent footprints are spatially

www.manaraa.com

localized and hence could suggest a sequential interaction. At 0°C the majority of S8 footprints are present, albeit with a significantly lower intensity. The exception is that three footprints in helix 26 are absent at low temperature (Figure 3a). On the structure of 16S rRNA from the 30S subunit, protections induced by S8 show a nice distribution along S8, and spanning from the body towards the head of the 30S subunit. The protections in the body appear at 0°C, along with some that continue to develop at 42°C which are closer to the head of the 30S subunit (Figure 4c). The protections that appear only at 42°C are located in the neck of the 30S subunit. It appears that S8 facilitates adjustments in the region towards the eventual head at higher temperature, after interacting with helix 21 across the back of the body at low temperature. Thus S8 might play a critical role in aligning the 3' major domain relative to other structural domains of the 30S subunit.

Once again our studies using the minimal S8/16S rRNA binary complexes reveal more details of this RNP, than assembly studies using a full complement of r-proteins. When binding of S8 is followed in the presence of all the r-proteins, during 30S assembly, no footprints specific for S8 were observed at 0°C.[17] Based on this data one cannot conclude if S8 binds to 16S rRNA at low temperature. However, when the minimal S8/16S rRNA complex is formed at the same temperature, the majority of S8 specific footprints are observed (Figure 3a). In fact, the majority of the footprints specific for S8 are observed at low temperatures, although many of them are only partial (Figure 3a and b). Thus our results are not in complete agreement with the classification of S8 as a mid-binding r-protein. The overall data might support a designation of early-mid

www.manaraa.com

binding protein for S8 since it can bind at an early phase, as epitomized by low temperature association, but than it is further accommodated later in assembly.

R-proteins that have distinct temperature dependent stages in the interaction with 16S rRNA. The head of the 30S subunit consists of short helical segments that are organized into a compact structure in the presence of many r-proteins. S7 is the only primary protein that binds to the 3' major domain initiating assembly,[3,15] and thus is critical for the binding of other r-proteins to this region. The area of 16S rRNA which is organized by S7 is very large, and this organization seems to require both S7 and a temperature-dependent conformational rearrangement. Our data suggest that binding of S7 occurs in two separable phases, allowing a model for the sequential interaction of S7 with 16S rRNA to be proposed. In this model, the highly charged N-terminus and the first helical elements of S7 would associate with 16S rRNA in the initial phase, while the second binding event involving the C-terminal portion of S7 would occur in the second phase, as revealed by temperature-dependent changes (Figure 5c, d). This idea is consistent with the studies in which the binding constants for the complexes formed between 16S rRNA and fragments of S7 were determined.[42] If the N-terminus of S7 is deleted binding to 16S rRNA is destroyed. When the N-terminus of S7 is intact but other parts of S7 are deleted, the binding constant decreases but the protein-RNA interaction still takes place.[42] These findings are consistent with *in vivo* studies which indicate that when the N-terminus of S7 r-protein is deleted, the assembly efficiency is reduced to about 3% of that observed for full length S7.[43] Thus our data reveal a model for a two stage association of S7 with 16S rRNA that is supported by other *in vitro* and *in vivo* studies, and likely reveal details of bipartite association of r-proteins with 16S rRNA.

www.manaraa.com

A temperature-dependent conformational rearrangement in the complex S4/16S rRNA was revealed by a chemical probing study,[27] and thus first revealed the utility of this approach for dissecting conformational rearrangements of 16S rRNA containing RNPs. In light of the results presented in this study, and with the advantage of information of the 30S subunit, discussion of the S4/16S rRNA complex should be revisited. The S4-specific protections that appear at 0°C are clustered near the lower part of S4 (Figure 5a), toward the bottom of 30S subunit structure. The protections which appear only at 42°C are located towards the head of 30S subunit, more proximal to the upper part of S4. From this, one could speculate that as suggested earlier for S7, the binding of S4 takes place in two steps. First, the central more globular domain of S4 would bind, and then the N- and C-termini would bind latter, as revealed by the temperature dependence. As the temperature-dependent S4 specific footprints are more dispersed throughout the 16S rRNA, and many of them are quite remote from the area of S4/16S rRNA direct contact (Figure 5b), it could be suggested that S4 facilitates long range conformational rearrangements during the course of 30S subunit assembly.

General trends in primary r-protein-16S rRNA interaction. An interesting correlation exists between temperature-dependent conformational rearrangements in the r-protein containing RNPs and the size of the r-protein (see Figure 6). RNPs containing the two smallest primary binding proteins S17 and S20 show no temperature dependence, while only a slight temperature dependence is observed for the RNPs containing the next smallest r-protein, S15. Continuing the trend, S8/16S rRNA shows more temperature dependence than the three smaller r-proteins mentioned above while RNPs containing S4 and S7, the largest primary binding r-proteins, show the highest degree of temperature-

www.manaraa.com

dependent conformational rearrangement. As expected, the smaller r-proteins tend to contain only one domain[3, 29] and may bind in a single event. The larger r-proteins tend to be composed from multiple "domains" or a more globular domain and some extended less canonical protein structures.[3,29] Thus these r-proteins may interact with 16S rRNA differentially.

The primary r-proteins S17 and S20 show little or no temperature dependence for the interaction with 16S rRNA. S20 and S17 are small, with highly regular protein structure.[3,29] NMR studies of S17 from *Bacilus stearothermophilus* (*B.st*.) showed that the core of the protein is in the same conformation as in the 30S subunit even in the free state[3, 10]. While, there are some differences in the structures of S17 from *E. coli* and *B. st*., the core structures are similar, with the main differences found in loops and tails. The core of the protein interacts with the 5' domain of 16S rRNA.[3,29] Interestingly, it was also shown that the 5' domain of naked 16S rRNA has structural features similar to those observed in the 30S subunit.[44] These two binding partners may be well structured in the unliganded form and associate in a single event. The structure of S20 in free form has not been determined yet, but in the 30S subunit S20 is a three helix bundle.[3,29] Since it also binds the 5' domain, and it has a very compact structure it is highly possible that it will behave similarly to S17. Thus, if both the rRNA and r-protein have structures similar to the bound forms before their interaction, it is expected that the conformational rearrangements at binding are going to be more minimal and thus the lack of temperature-dependent conformational rearrangements is not surprising.

For S15, the situation is different than for S17 and S20. Structural studies suggest that S15 does not undergo large conformational changes upon binding, although the N-

www.manaraa.com

terminal helix of S15 may be dynamic and thus change orientation.[2,11,37] Conversely, it has been demonstrated that 16S rRNA goes through a big conformational rearrangement upon S15 binding.[2] Our results suggest that the conformational changes in rRNA can largely take place even in the absence of heating (Figure 3). Thus, for the S15/16S rRNA complex, induced fit likely occurs mainly in a single event and might be largely restricted to the rRNA.

For the RNPs containing r-proteins that have multiple domains (S4, S7 and S8) more temperature-dependent conformational rearrangements are observed. Beside the fact that they have more complex structures, their binding also affects more extended regions of the 16S rRNA. Therefore, it is likely that these r-proteins play a role in organizing long range interactions during the 16S rRNA folding and 30S subunit assembly. For S7, it was shown that the structure of the free protein and the protein in 30S subunit are different,[3,7,8] while for S4 it seems that the core structure is the same before and after binding[3,9,12,29] although information is only available for a subfragment of S4 in the free form. In the case of S8, whose binding is important for the organization of the central domain, temperature augments the number of footprints with 16S rRNA. For S4 and S7, primary r-proteins whose binding shapes large regions of 16S rRNA, temperature plays a very important role in the conformation of the RNP containing either r-protein. S4 and S7 show distinct temperature dependent stages and in general, the long distance effects are observed at higher temperature. Major conformational rearrangements in the rRNA are taking place upon heating the complex, and allow further events in the assembly of the 30S subunit.

www.manaraa.com

Implications for 30S subunit assembly: Assembly of secondary r-proteins. The conformational changes observed for some of the 16S rRNA-primary r-protein containing RNPs have significant implications for the subsequent assembly events. The majority of the experiments performed to analyze the assembly dependence of the small subunit r-proteins were carried out at elevated temperature. Thus the work presented here aids in our understanding of the requisite order that has been observed.

In the important 530 loop, enhancement of specific nucleotides requires both S8[25] and elevated temperature. In subsequent stages of assembly these enhanced nucleotides become less reactive to chemical probes in an S5-dependent manner.[45] Thus r-protein mediated temperature-dependent conformational changes can be critical for perpetuation of the 30S subunit assembly cascade.

One of the secondary binding r-proteins that has a very important role in the formation of the first low temperature intermediate in the assembly of the 30S subunit (RI) is S16.[20,21,46] Chemical probing experiments have been performed to determine changes in 16S rRNA folding during the transitions from naked 16S rRNA to RI, RI to RI* and finally RI* to fully assembled 30S subunits, using mixtures of r-proteins.[20,21] These experiments attribute 13 changes in reactivity observed only in the 16S rRNA to RI transition to S16 alone or S16 in combination with S20 (seven changes).[20,21] Thus S16 must be able to bind to a 16S rRNA containing RNP at low temperature. Interestingly, S20 and S4 are the two primary proteins required for the assembly of S16 (Figure 1a).[47] The S20/16S rRNA complex does not seem to undergo a temperature-dependent conformational rearrangement and thus once S20 has bound at low temperature it is possible that S16 can associate. It was shown that binding of S16 to the region formed by

www.manaraa.com

the nucleotides 1-353 and the penultimate stem is mainly controlled by the initial binding of S20.[35] Binding of S16 is also dependent on the interaction of S4 with 16S rRNA, as it was mentioned earlier. S16 has many footprints both in the 5' domain and the central domain.[35] Probably, S4 modulates binding of S16 to other regions of 16S rRNA, in particular the central domain. It is possible that S4 and S16 containing RNPs undergo synergistic temperature-dependent changes[20,21] that are then important for assembly of additional r-proteins.

The secondary binding proteins which depend on the initial assembly of S15 are S18 and S6.[47] However S18 and S6 (along with S15) are also classified as early binders.[17] The enhancements produced by the binding of S15 (G664, G674, C719) that become protected by binding of S6 and S18[25] are not temperature dependent (Figure 3). Consequently this would favor the rapid binding of the secondary proteins, even at low temperature as would be consistent with their kinetic classification. Again, these events may then allow assembly in an ordered manner.

Binding of S7 prepares 16S rRNA for the binding of other r-proteins like the secondary binding r-protein S19.[15] The regions around 950 and 980 reveal temperature – dependent changes (Figure 3c), mainly enhancements, in the S7/16S rRNA RNP, although there are no crystal contacts between 16S rRNA and S7 in these regions in the 30S subunit.[3,29] The temperature-dependent enhancements are sites that can be protected by binding of S19.[26] Thus the temperature-dependent conformational rearrangement of S7/16S rRNA complex likely facilitates full accommodation of S19. This proposal is consistent with the classification of S19 as a mid-binding protein.[17]

www.manaraa.com

Temperature-dependent conformational rearrangements may be a common device in orchestrating an orderly sequential assembly in complex RNPs that involve large RNAs. Binding of the r-proteins facilitates changes in 16S rRNA conformation at the RNA-protein interface, but it was also shown that it affects the conformation of 16S rRNA at some distance from the site of interaction. These long distance effects can organize the binding site of other r-proteins that assemble in a sequential manner, modulate subunit interdomain interactions or bring the 16S rRNA into a correct functional conformation. For some of these long distance effects to be realized, the appropriate changes in the conformation of 16S rRNA require both the presence of an r-protein and elevated temperature. Thus the approach may be applicable to the study of large RNP assembly.

The complexity of the spectrum of interactions between RNA and proteins is very well illustrated in our model system. Even in the same RNP, differential folding of segments of the RNA molecule with a single protein can be observed. Very simple single phase interactions are observed, in general, with RNA binding proteins that are very small, and well structured. In other instances, a more regulated process appears to be utilized. Alteration of temperature can be used to reveal modulation of folding of RNA within these RNP. Changes in RNA structure within the RNP can be subtle, such as fine adjustments, or quite substantial. Our simple study suggests that the r-proteins can interact with 16S rRNA differentially and that at least two types of induced fit are observed: when only RNA is changing its conformation after binding and when both the RNA and the protein are changing conformation at binding.

www.manaraa.com

References

1. Leulliot, N.; Varani, G., *Biochemistry* **2001,** 40, (27), 7947-56.
2. Williamson, J. R., *Nat. Struct. Biol.* **2000,** 7, (10), 834-7.
3. Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H., *Science* **2005,** 310, (5749), 827-34.
4. Lynch, S. R.; Gonzalez, R. L.; Puglisi, J. D., *Structure* **2003,** 11, (1), 43-53.
5. Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V., *Nature* **2000,** 407, (6802), 340-8.
6. Wimberly, B. T.; Brodersen, D. E.; Clemons, W. M., Jr.; Morgan-Warren, R. J.; Carter, A. P.; Vonrhein, C.; Hartsch, T.; Ramakrishnan, V., *Nature* **2000,** 407, (6802), 327-39.
7. Hosaka, H.; Nakagawa, A.; Tanaka, I.; Harada, N.; Sano, K.; Kimura, M.; Yao, M.; Wakatsuki, S., *Structure* **1997,** 5, (9), 1199-1208.
8. Wimberly, B. T.; White, S. W.; Ramakrishnan, V., *Structure* **1997,** 5, (9), 1187-98.
9. Sayers, E. W.; Gerstner, R. B.; Draper, D. E.; Torchia, D. A., *Biochemistry* **2000,** 39, (44), 13602-13613.
10. Golden, B. L.; Hoffman, D. W.; Ramakrishnan, V.; White, S. W., *Biochemistry* **1993,** 32, (47), 12812-20.
11. Berglund, H.; Rak, A.; Serganov, A.; Garber, M.; Hard, T., *Nat. Struct. Biol.* **1997,** 4, (1), 20-3.
12. Davies, C.; Gerstner, R. B.; Draper, D. E.; Ramakrishnan, V.; White, S. W., *EMBO J.* **1998,** 17, (16), 4545-58.

www.manaraa.com

13. Traub, P.; Nomura, M., *Proc Natl Acad Sci U S A* **1968,** 59, (3), 777-84.

14. Noller, H. F.; Woese, C. R., *Science* **1981,** 212, (4493), 403-11.

15. Mizushima, S.; Nomura, M., *Nature* **1970,** 226, (5252), 1214-18.

16. Grondek, J. F.; Culver, G. M., *RNA* **2004,** 10, (12), 1861-1866.

17. Powers, T.; Daubresse, G.; Noller, H. F., *J. Mol. Biol.* **1993,** 232, (2), 362-74.

18. Traub, P.; Nomura, M., *J. Mol. Biol.* **1969,** 40, (3), 391-413.

19. Held, W. A.; Nomura, M., *Biochemistry* **1973,** 12, (17), 3273-81.

20. Holmes, K. L.; Culver, G. M., *J. Mol. Biol.* **2005,** 354, (2), 340-357.

21. Holmes, K. L.; Culver, G. M., *Nat. Struct. Biol.* **2004,** 11, (2), 179-186.

22. Guthrie, C.; Nashimoto, H.; Nomura, M., *Proc. Natl. Acad. Sci. U.S.A.* **1969,** 63, (2), 384-91.

23. Nashimoto, H.; Held, W.; Kaltschmidt, E.; Nomura, M., *J. Mol. Biol.* **1971,** 62, (1), 121-38.

24. Nierhaus, K. H.; Bordasch, K.; Homann, H. E., *J. Mol. Biol.* **1973,** 74, (4), 587-97.

25. Svensson, P.; Changchien, L. M.; Craven, G. R.; Noller, H. F., *J. Mol. Biol.* **1988,** 200, (2), 301-8.

26. Powers, T.; Changchien, L. M.; Craven, G. R.; Noller, H. F., *J. Mol. Biol.* **1988,** 200, (2), 309-19.

27. Powers, T.; Noller, H. F., *J. Biol. Chem.* **1995,** 270, (3), 1238-42.

28. Nowotny, V.; Nierhaus, K. H., *Biochemistry* **1988,** 27, (18), 7051-5.

29. Brodersen, D. E.; Clemons, W. M., Jr.; Carter, A. P.; Wimberly, B. T.; Ramakrishnan, V., *J. Mol. Biol.* **2002,** 316, (3), 725-68.

www.manaraa.com

30. Baumann, C.; Otridge, J.; Gollnick, P., *J. Biol. Chem.* **1996,** 271, (21), 12269-74.

31. Moazed, D.; Stern, S.; Noller, H. F., *J. Mol. Biol.* **1986,** 187, (3), 399-416.

32. Culver, G. M.; Noller, H. F., *Methods Enzymol.* **2000,** 318, (RNA-Ligand Interactions, Pt. B), 446-460.

33. Culver, G. M.; Noller, H. F., *RNA* **1999,** 5, (6), 832-43.

34. Merryman, C.; Noller, H. F., Footprinting and modification-interference analysis of binding sites on RNA. In *RNA:Protein Interactions. A Practical Approach*, Smith, C. W. J., Ed. Oxford University Press: New York, 1998.

35. Stern, S.; Changchien, L. M.; Craven, G. R.; Noller, H. F., *J. Mol. Biol.* **1988,** 200, (2), 291-9.

36. Yusupov, M. M.; Yusupova, G. Z.; Baucom, A.; Lieberman, K.; Earnest, T. N.; Cate, J. H.; Noller, H. F., *Science* **2001,** 292, (5518), 883-96.

37. Agalarov, S. C.; Sridhar Prasad, G.; Funke, P. M.; Stout, C. D.; Williamson, J. R., *Science* **2000,** 288, (5463), 107-13.

38. Gerstner, R. B.; Pak, Y.; Draper, D. E., *Biochemistry* **2001,** 40, (24), 7165-73.

39. Mougel, M.; Eyermann, F.; Westhof, E.; Romby, P.; Expert-Bezancon, A.; Ebel, J. P.; Ehresmann, B.; Ehresmann, C., *J. Mol. Biol.* **1987,** 198, (1), 91-107.

40. Mougel, M.; Ehresmann, B.; Ehresmann, C., *Biochemistry* **1986,** 25, (10), 2756-65.

41. Talkington, M. W.; Siuzdak, G.; Williamson, J. R., *Nature* **2005,** 438, (7068), 628-32.

42. Robert, F.; Gagnon, M.; Sans, D.; Michnick, S.; Brakier-Gingras, L., *RNA* **2000,** 6, (11), 1649-59.

www.manaraa.com

43. Fredrick, K.; Dunny, G. M.; Noller, H. F., *J. Mol. Biol.* **2000,** 298, (3), 379-394.

44. Adilakshmi, T.; Ramaswamy, P.; Woodson, S. A., *J. Mol. Biol.* **2005,** 351, (3), 508-19.

45. Stern, S.; Powers, T.; Changchien, L. M.; Noller, H. F., *J. Mol. Biol.* **1988,** 201, (4), 683-95.

46. Held, W. A.; Nomura, M., *J. Biol. Chem.* **1975,** 250, (8), 3179-84.

47. Held, W. A.; Ballou, B.; Mizushima, S.; Nomura, M., *J. Biol. Chem.* **1974,** 249, (10), 3103-11.

48. DeLano, W. L., *DeLano Scientific* **2002**, San Carlos, CA, USA.

49. Cannone, J. J.; Subramanian, S.; Schnare, M. N.; Collett, J. R.; D'Souza, L. M.; Du, Y.; Feng, B.; Lin, N.; Madabusi, L. V.; Muller, K. M.; Pande, N.; Shang, Z.; Yu, N.; Gutell, R. R., *BMC Bioinformatics* **2002,** 3, 2.

www.manaraa.com

Table 1. 16S rRNA nucleotides with changed reactivity as a result of r-protein binding.

The type of change is indicated and if the change is temperature dependent or not.

	S20		S17		S8		S15		S7	
Nucleotide	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.
5` domain										
A182	P	No								
A189	E	No								
A190	E	No								
C194	P	No								
A246	P	No	P	No						
A250	P	No	P	No						
G251	P	No	P	No						
A262	P	No	P	No						
A263	P	No	P	No						
C264	P	No	P	No						
G265	P	No	P	No						
G266	P	No	P	No						
A274	P	No	P	No						
A279	P	No	P	No						
C280	P	No	P	No						
G281	P	No	P	No						
A325	E	No								
A327	P	No								
A329	P	No								
G331	P	No								
G332	P	No								
A338	E	No								
C352	P	No								
A353	P	No								
G524					E	Yes	E	Yes		
A535					E	Yes	E	Yes		
Central domain										
A573					P	Yes				
A574					P	Yes				
G575					P	Yes				
A583					P	No				
A640					P	No				
A642					P	No				
G664							E	No		
A665							E	Yes		

www.manaraa.com

Nucleotide	S20		S17		S8		S15		S7	
	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.
A673					E	No	E	No		
G674					E	Yes	E	No		
A718					E	Yes	E	Yes		
G724							P	Yes		
G727							P	Yes		
A728					E*	Yes	E*	Yes		
G730							P	No		
C732					E*	Yes				
G741							P	No		
G742							E*	No		
C754					P*	No				
G812					P	No				
C817					E	No				
G858					P	Yes				
G859					P	Yes				
A860					P	Yes				
G861					P	Yes				
A865					P	Yes				
3`major domain										
A935									P	Yes
C936									P	No
A937									P	Yes
A938									P	Yes
G939									P	Yes
C940									P	No
G944									P	Yes
G945									P	Yes
A949									P	No
G951									P	Yes
G953									E *	Yes
G954									E	Yes
A977									E	Yes
A978									E	Yes
C979									E	Yes
C980									E	Yes
A983									P	No
A1236									P*	Yes
C1237									P*	Yes
A1238									P	Yes
A1239									P*	Yes
A1248									P	Yes

www.manaraa.com

Nucleotide	S20		S17		S8		S15		S7	
	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.
C1249									P	Yes
A1250									P	Yes
A1251									P	Yes
A1252									P	Yes
A1256									P	No
A1287									P	Yes
A1288									P	Yes
A1289									P	Yes
G1290									P	No
G1297									E	No
G1300									P	Yes
C1302									P	Yes
G1304									P	Yes
G1305									P	Yes
C1314									P*	Yes
G1316									P	Yes
C1317									E*	Yes
A1318									E*	Yes
A1319									E*	Yes
C1320									E*	Yes
C1322									P	Yes
G1331									P	Yes
A1332									P	Yes
A1333									P	Yes
G1334									P	Yes
G1337									E *	Yes
G1338									E	Yes
A1339									P	No
A1346									P	No
A1349									P	Yes
A1360									P	Yes
G1361									P	Yes
A1362									P	Yes
A1363									P	Yes
G1365									P	Yes
A1374									P	No
A1377									P	No
C1382									P	No
3’ minor domain										
A1433	P	No								
A1434	P	No								

www.manaraa.com

Nucleotide	S20		S17		S8		S15		S7	
	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.	Type	T^0C dep.
A1446	P	No								
A1447	P	No								
C1469	E	No								

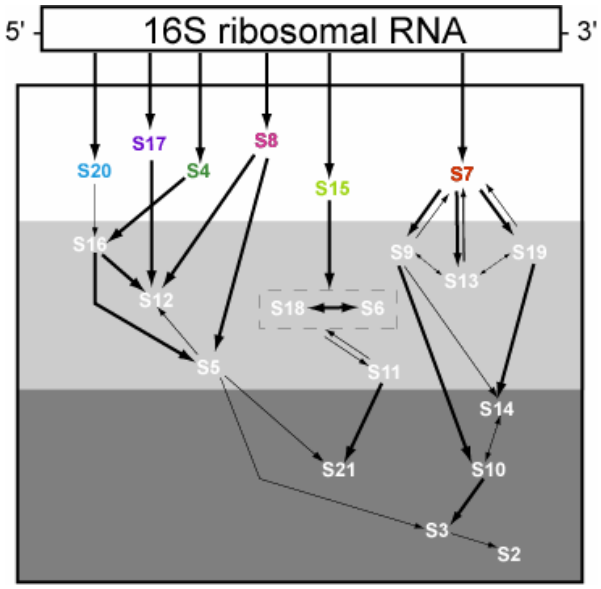
E –enhancement, P-protection.

* -indicates that no change in reactivity was reported previously at that nucleotide

www.manaraa.com

Figure 1. (a) Modified *in vitro* 30S subunit assembly map. The 16S rRNA is represented by a rectangle in a 5' to 3' direction. The arrows indicate the co-dependencies for the assembly of the r-proteins. The relative size of the arrow indicates the relative strength of the assembly dependency between components. The r-proteins shown in the white region are primary r-proteins. The r-proteins shown in white in the light gray and dark gray box indicate secondary, and tertiary binding r-proteins, respectively. S6 and S18 are enclosed in a box to indicate that they bind as a heterodimer. (b) Crystal structure of the 16S rRNA from the *E. coli* 30S subunit with all the primary proteins.[3] The 16S rRNA is shown in gray, and the r-proteins are S4 green, S7 red, S8 magenta, S15 bright yellow, S17 dark purple and S20 light blue. All the Figures containing 3-D structures were prepared using Pymol,[48] and the pdb file 2AW7.

www.manaraa.com



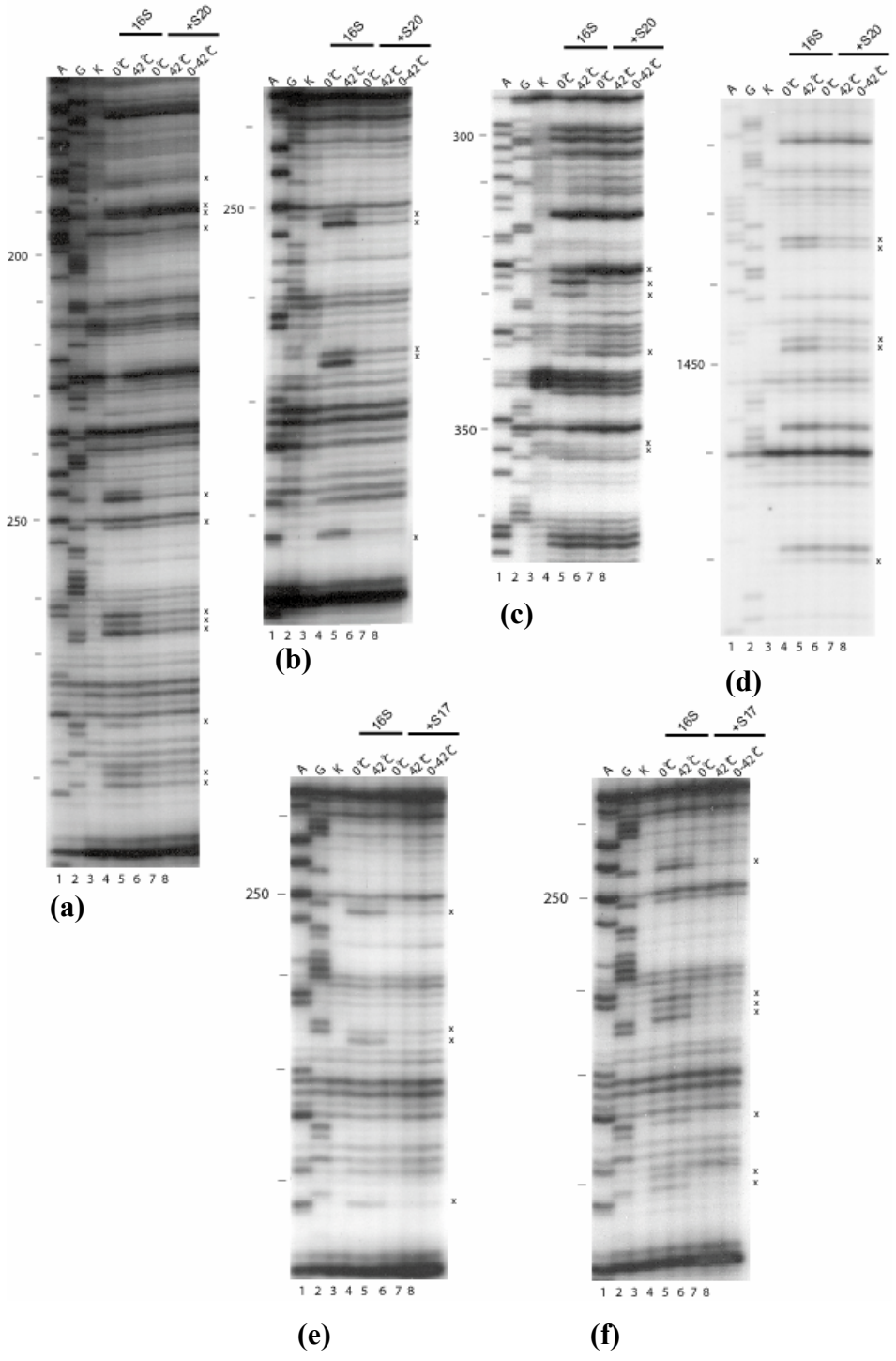
(a)

(b)

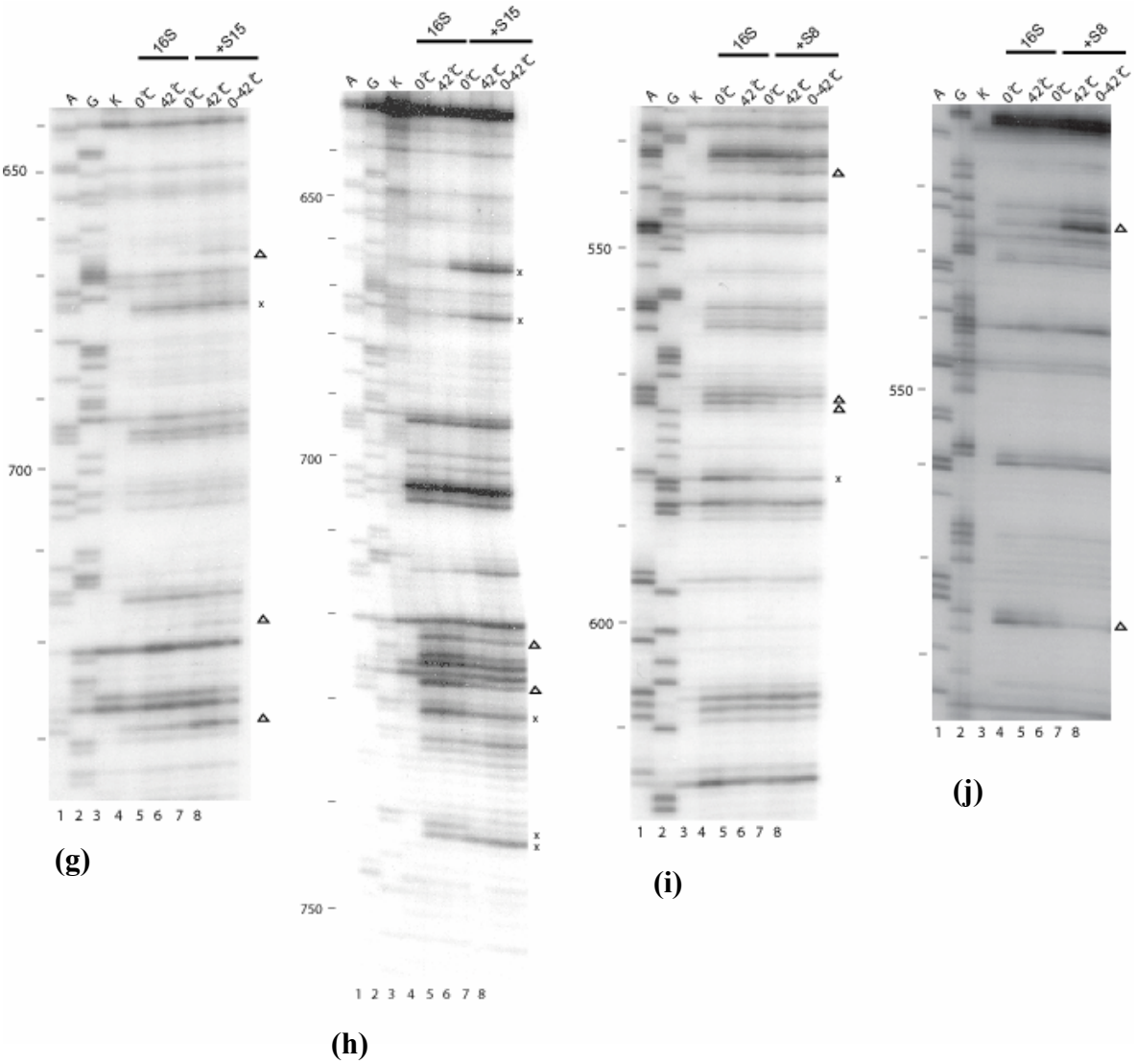
www.manaraa.com

Figure 2. Primer extension analysis of the r-protein/16S rRNA complexes. Individual gels of the minimal complexes modified by DMS or kethoxal are shown. A and G (lanes 1 and 2) are dideoxy sequencing lanes, K (lane 3): unmodified 16S rRNA. All the other lanes are treated with the probe indicated below. The other lanes 4-8 are: modified 16S rRNA kept at 0^{o}C (lane 4) at 42^{o}C (lane 5), Sx/16S rRNA formed at 0^{o}C (lane 6), at 42^{o}C (lane 7) and the shifted complex (lane 8). Compare lanes 4 and 6 for the complexes formed at 0^{o}C, lanes 5 and 7 for the complexes formed at 42 ^{o}C, or lanes 6 and 7 for the differences between the two complexes. The primers used for the extension are indicated below. S20/16S rRNA: (a) DMS-323, (b) kethoxal-323, (c) DMS-480, (d) DMS-1508; S17/16S rRNA: (e) DMS -323; (f) - kethoxal, 323; S15/16S rRNA: (g) DMS - 795 , (h) kethoxal – 795; S8/16S rRNA: (i) DMS-683, (j) kethoxal-683, (k) DMS-795, (l) DMS-939, (m) kethoxal -939 , (n) DMS – 939; S7/16S rRNA: (o) DMS-1046, (p) kethoxal -1046, (q) DMS-1391, (r) DMS -1491, (s) kethoxal -1491. The symbol x indicates no temperature dependence and the symbol Δ indicates temperature dependence.

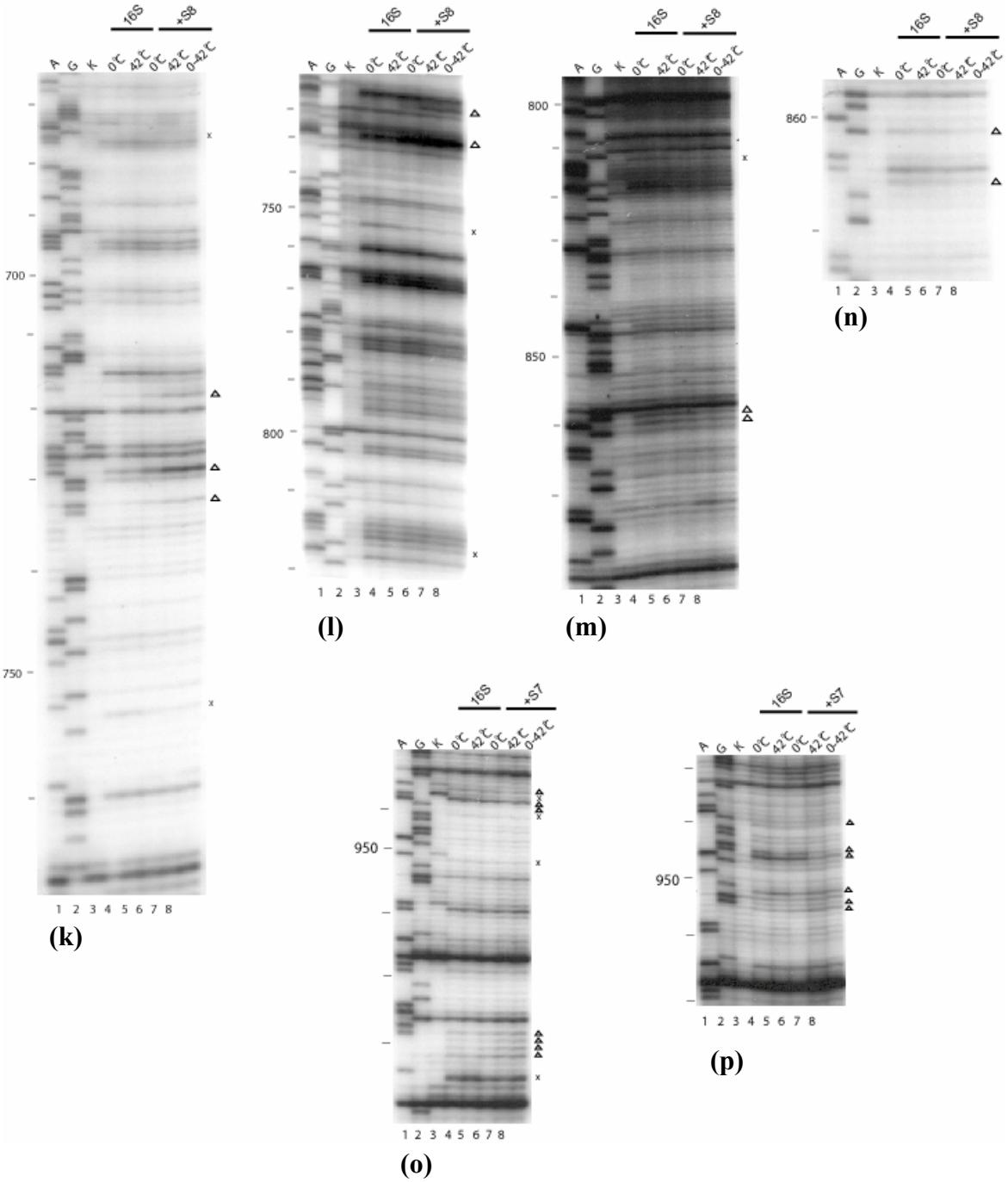
www.manaraa.com



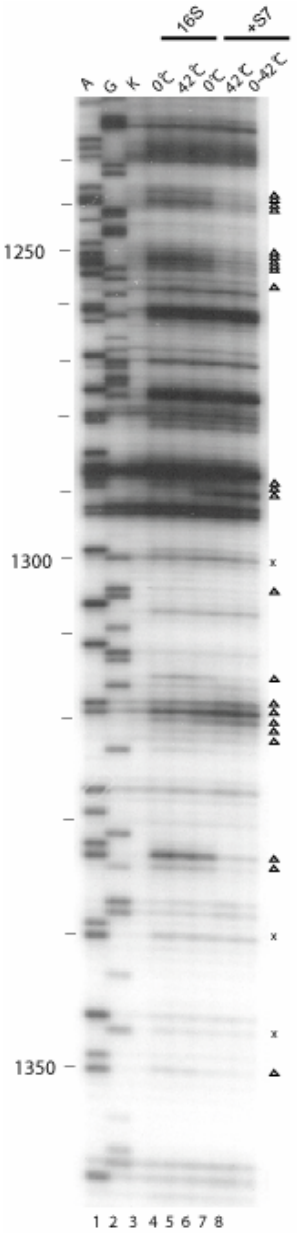
www.manaraa.com



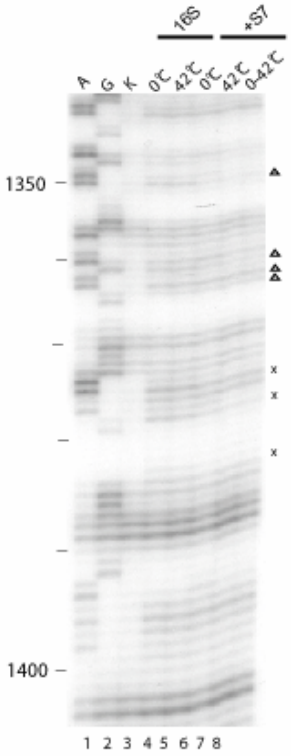
www.manaraa.com



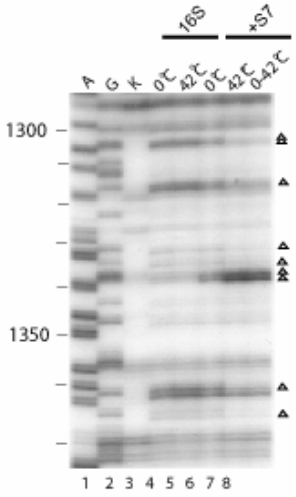
www.manaraa.com



(q)



(r)

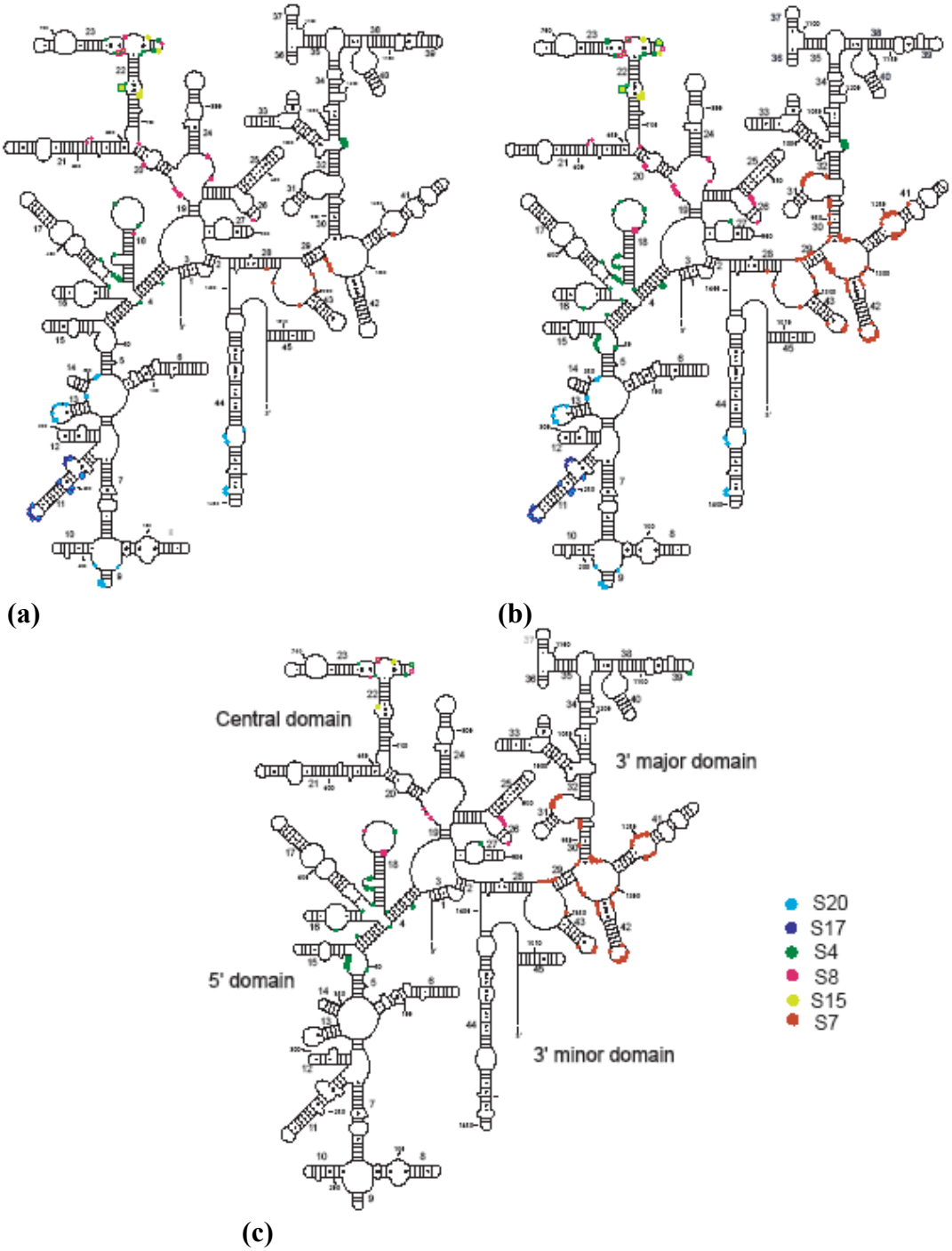


(s)

www.manaraa.com

Figure 3. Nucleotides with altered reactivity as a result of binding of an r-protein to 16S rRNA, for both kethoxal and DMS probing represented on the secondary structure of 16S rRNA.[49] Circles denote the sites of protections, while squares denote enhancement sites, and the size represents the intensity of the change. The 16S rRNA is shown in gray, changes attributed to: S4, green; S7, red; S8, magenta; S15 bright yellow; S17, dark purple, and S20, light blue. Nucleotides enhanced or protected by more than one protein are shown as concentric rings or squares. (a) Changes in modification patterns shown on the secondary structure of 16S rRNA for the interaction at 0^{o}C. (b) Changes in modification patterns shown on the secondary structure of 16S rRNA for the interaction at 42^{o}C. (c) Difference in the nucleotides with altered reactivity between the complexes formed at 42^{o}C and 0^{o}C.

www.manaraa.com

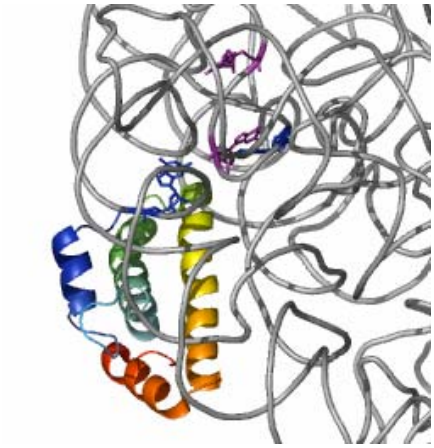


www.manaraa.com

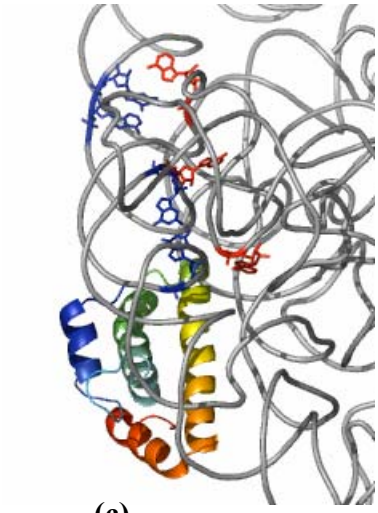
Figure 4. Details of footprints for r-proteins S15 and S8. The 16S rRNA is shown in gray, S15 and S8 are shown in rainbow, from blue (N-terminus) to red (C-terminus). Footprints which are not temperature dependent are shown in blue, footprints that appear at 0^{o}C and continue to develop in intensity at 42^{o}C are shown in purple, and footprints that appear only at 42^{o}C are shown in red. (a) 16S rRNA r-protein and S15 from the crystal structure of E. coli 30S subunit; (b) S15-dependent protections; (c) S15-dependent enhancements; (d) 16S rRNA r-protein and S8 from the crystal structure of E. coli 30S subunit; (e)S8-dependent protections; (f) S8-dependent enhancements.

www.manaraa.com

(a)

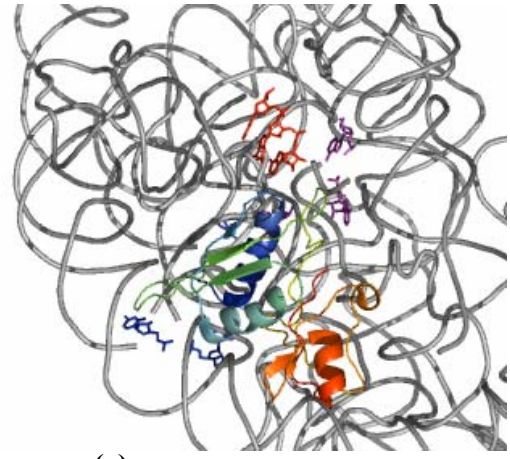


(b)

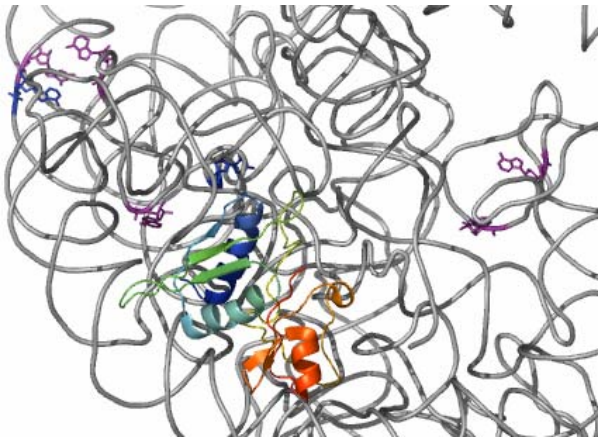


(c)

(d)



(e)



(f)

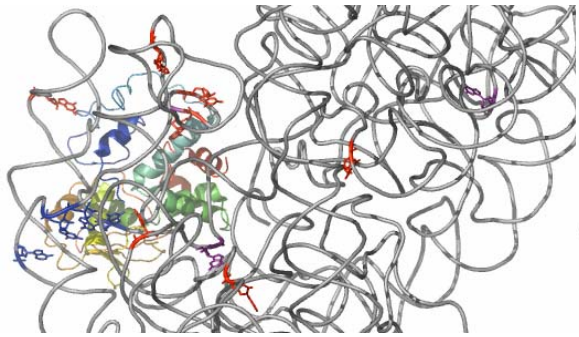
www.manaraa.com

Figure 5. Details of footprints for r-proteins S4 and S7. The 16S rRNA is shown in gray, S15 and S8 are shown in rainbow, from blue (N-terminus) to red (C-terminus). Coloring of footprints as described in Figure 4. (a) 16S rRNA r-protein and S4 from the crystal structure of *E. coli* 30S subunit; (b) S4-dependent protections; (c) S4-dependent enhancements; (d) 16S rRNA r-protein and S7 from the crystal structure of *E. coli* 30S subunit; (e) S7-dependent protections; (f) S7-dependent enhancements.

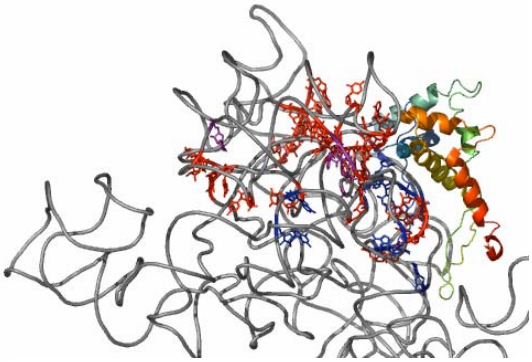
www.manaraa.com

(a)

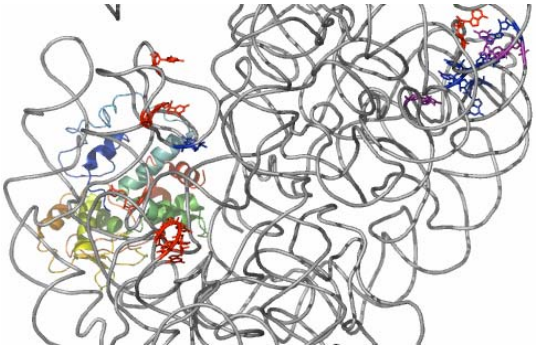
(d)



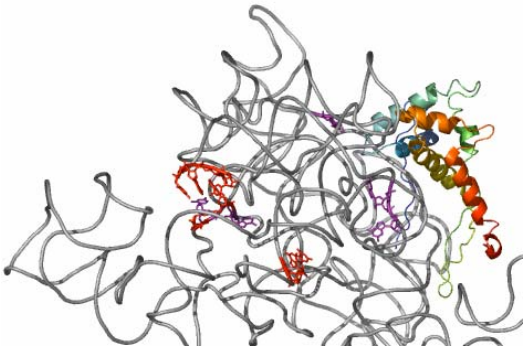
(b)



(e)



(c)

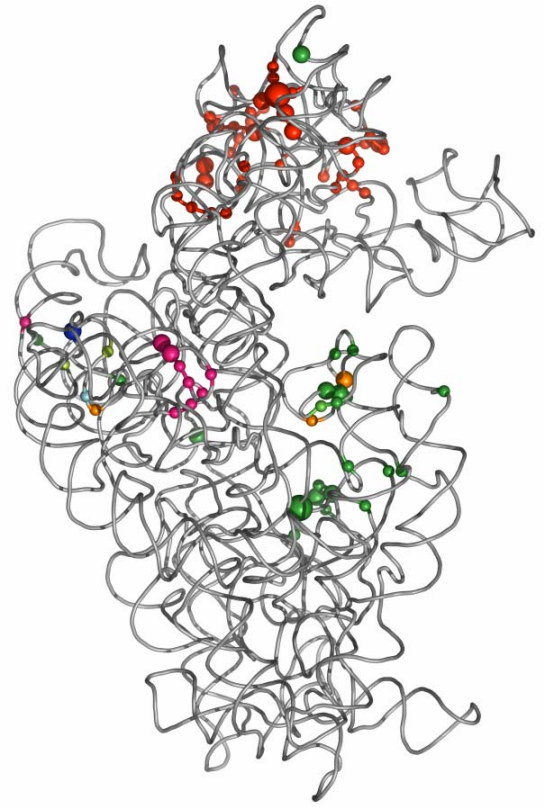


(f)

www.manaraa.com

Figure 6. The temperature dependent footprints for each primary r-protein represented on the crystal structure of 16S rRNA from the *E. coli* 30S subunit, as spheres. The proteins are not shown for clarity. The size of the spheres is indicative of the intensity of the change. The 16s rRNA is shown in gray. The temperature dependent footprints (spheres) correspond to: S4, green; S7, red; S8, purple; S15, bright yellow; S4 and S15, magenta; S8 and S15 orange; S4, S8 and S15, blue. Different views of 16S rRNA from 30S subunit are shown: (a) solvent view and (b) interface view.

www.manaraa.com



(a)

(b)

www.manaraa.com

Chapter VII. Probing the assembly of 30S subunit with the r-protein S20

Abstract

The ribosomal protein (r-protein) S20 interacts directly and independently with the 5' domain and the 3' minor domain of 16S ribosomal RNA (rRNA) in minimal particles and the fully assembled 30S subunit. In this study Fe(II)-derivatized S20 protein is used as a probe for the assembly of the 30S subunit. Directed hydroxyl radical probing from four unique positions on S20 reveals the architecture of 16S rRNA around the probe. An analysis of the cleavage patterns in the minimal complexes and the fully assembled 30S subunit shows intriguing similarities and differences. For a better understanding of the events taking place during assembly around the probe, Fe(II)-S20, ribonucleoprotein particles (RNPs) of different complexities are probed. The comparison of the cleavage patterns in the different RNPs shows that even in the minimal particles the environment of the probe is very similar to the one in the 30S subunit, but addition of other r-proteins augments the organization of 16S rRNA.

Introduction

The process by which 16S ribosomal RNA (rRNA) folds into three-dimensional structures within functional 30S ribosomal subunits has drawn much interest, but there are still many unanswered questions. The crystal structure of the ribosome from *Escherichia coli* (*E. coli*) was determined recently,[1] and structures of the ribosome and the independent subunits are available from other organisms[2-5] too. The knowledge of the structure of some of the ribosomal proteins[6-10] (r-proteins) and fragments of the rRNA[11]

www.manaraa.com

also available in free form make possible some inferences on the assembly process. These advances are very important but they give us information mostly about the static complexes without revealing the dynamics of RNA and r-proteins during assembly. The next challenge in understanding 30S subunit assembly is to identify changes the changes that lead to the formation of a mature, functional 30S subunit. The complex composition of the 30S ribosomal subunit, 21 r-proteins (S1-S21) and 16S rRNA,[12] makes this problem non-trivial and only by using a variety of methods will the processes involved in ribosome biogenesis be understood.

16S rRNA has 1542 nucleotides, and studies revealed four distinct secondary structural domains, the 5', central, 3' major and 3' minor domains[13,14] (Figure 1a). The four domains of the 16S rRNA form, in the presence of the appropriate r-proteins form distinct parts of the 30S subunit, that can assemble independently, and they are the body (5' domain), the platform (central domain), the head (3' major domain) and the penultimate stem (mainly the 3' minor domain).[1,13] The r-proteins are the driving force behind the assembly of the 30S subunit. Their binding shapes and defines the structure of the 16S rRNA in the functional small subunit. The assembly map[15,16] (Figure 1b) is a guide in the study of the assembly of the 30S subunit. *In vitro* experiments that used the self-assembly capacity of the 30S subunit, in which single r-proteins where omitted or added revealed the requirements for and the order of r-protein binding to 16S rRNA.[15,17,18] The small subunit r-proteins are classified based on the requirements for their assembly in three categories: primary, secondary and tertiary[15,17,18] (Figure 1b). The primary binding r-proteins bind independently and specifically to 16S rRNA, in the absence of any other r-proteins. The secondary and tertiary binding r-proteins necessitate

www.manaraa.com

the prior assembly of at least one primary r-protein, and of both primary and secondary binding r-proteins, respectively.

The 30S subunit can assemble *in vitro* from 16S rRNA and a mixture of r-proteins,[19] which can be extracted from the natural 30S subunits, as a mixture,[12] separated individually or from recombinant r-proteins.[20] The recombinant system for *in vitro* reconstitution makes possible the construction of minimal ribonucleoprotein particles (RNPs) that represent different stages of assembly, in which r-proteins that are modified or derivatized can be incorporated.

One of the methods used extensively, but mainly in static complexes, to study the nucleic acid environment of proteins, or even nucleic acids is directed hydroxyl radical probing.[21,22] Proteins are derivatized with Fe(II)BABE through a single cysteine,[21] and RNA through phosphorothioate.[23] In directed hydroxyl radical probing, the radicals are generated only around the tethered Fe(II) by Fenton reaction, cleaving the RNA backbone[21]. In the ribosome and ribosomal subunits it helped in the characterization of the rRNA surroundings of components such as r-proteins[21,24,25] or ligands like transfer RNA,[23,26,27] prior to structural advances. For example, the location of r-protein S20 in 30S subunit was analyzed, using derivatized S20 as a probe, and helped clarify the controversy of its location[21]. Recently, directed hydroxyl probing was used to explore the dynamics of the 16S rRNA surrounding the r-protein S15.[28,29] Derivatized S15 was used to identify conformational changes in the 16S rRNA in RNPs of different complexities,[28,29] thus elucidating roles of the r-proteins in the assembly process. Fe(II)-tethered S15 protein was incorporated in different complexes and the changes in the cleavage patterns were used to asses the changes in the rRNA structure. The minimal

www.manaraa.com

complex (Fe(II)-S15/16S rRNA) showed the least complex cleavage pattern, and at the addition of other r-proteins, more cleavage sites were observed.[29] The difference in the cleavage pattern between the complexes shed light on the rearrangement of the rRNA elements in the presence of a certain r-protein.[28] This approach made possible a better understanding of the assembly of the 30S subunit, by detailing the role played by r-proteins and the conformational rearrangements that occur in nucleic acid-protein complexes during assembly.

Some of the 30S subunit r-proteins have more than one 16S rRNA binding site.[4] It has been suggested that the interaction between r-proteins and 16S rRNA occurs in discrete stages.[30-32] Early in 30S subunit assembly the protein could interact with a specific 16S rRNA element and later interaction with a second site of 16S rRNA could occur[31]. One of the ribosomal proteins that interacts with two different domains of 16S rRNA is a primary binding protein S20[4,33](Figure 2a and b). The interaction of primary r-protein S20 with 16S rRNA has been studied extensively before the structure of the 30S subunit was determined. Footprinting experiments with base-specific chemical probes and solution hydroxyl radical probing of the minimal complex (S20/16S rRNA) were used to reveal the binding site to the 16S rRNA.[34] The crystal structure of the 30S subunit confirmed that S20 binds several helices from the 5'domain and the 3' minor domain of 16S rRNA.[1] Its structure was not determined in the free form, but in the assembled 30S subunit, S20 is a three-helix bundle located at the bottom of the small subunit in the body (5' domain), and it also contacts the penultimate stem (helix 44)[1] (Figure 2a and b). Based on the model of 5' to 3' assembly[32] and the positioning of the penultimate stem across the body in the small subunit,[1] it is easy to speculate that S20 may interact with

www.manaraa.com

these domains differently along the assembly, and directed hydroxyl radical probing may be a suitable method to dissect these interactions.

The presence of cysteine residues is required for directed hydroxyl radical probing, but the wild-type r-protein S20 does not contain any cysteine residues. Since directed hydroxyl radical probing from S20 in the 30S subunit was effected earlier,[33] the positions at which individual cysteines could be introduced were already available.[33] Non-conserved residues chosen using an amino acid sequence alignment of S20 proteins from five organisms, which had a high probability to be found on the surface of the protein were replaced by cysteines. Four modified proteins that gave base specific footprints similar to those of the wild type S20 and thus retained function, were used for directed hydroxyl radical probing.[33] At that time, the crystal structure of S20 was not known, and the placement of cysteines on the tertiary structure of S20 could not be determined, but they gave different cleavage patterns, so it was presumed that they are fairly well dispersed along S20. In the crystal structure of S20 from the 30S subunit (Figure 2e, f) one of the cysteines is close to the N-terminus of S20, at one end of S20 (residue 13), another one is at the opposite end of S20 (residue 47), while the remaining two (residues 22 and 55) are in the middle of S20, very close in space, but on different helices (Figure 2e, f). In the experiments we have used the four aforementioned substituted S20 proteins: C13S20, C22S20, C47S20 and C55S20.

Initially, the minimal Fe(II)–S20 /16S rRNA complexes are analyzed and their cleavage patterns are compared to the ones in the 30S subunits,[33] both from literature and reproduced. The cleavage patterns in the minimal complexes and the fully assembled 30S subunits are very similar, which is very different from what was observed for r-protein

www.manaraa.com

S15.[28,29] Traditionally, particles with the same composition but containing one r-protein derivatized at individual distinct positions were probed. This approach is useful if there is not a lot of information on the structure of the complex formed between the r-protein and the 30S subunit, since it will provide a lot of information on the surroundings of the probe. In our case, because the cleavage patterns that we obtained for the starting complex (Fe(II)–S20 /16S rRNA) were very similar to the ones observed for the final product of assembly (30S subunit), we thought that it would be more informative to probe at the same time RNPs of different complexities that contain the same derivatized S20 protein. This approach should confirm the similarities and reveal subtle differences. In this study our results for the probing of minimal complexes are presented, followed by the results for the probing of the 30S subunit, and to obtain a more detailed picture of the dynamics of the surroundings of S20 during assembly, different RNPs containing derivatized S20 protein, in particular C13S20 and C22S20, were probed in the same time and the results are analyzed.

Materials and methods

Mutagenesis, expression and purification of S20. The gene-encoding ribosomal protein S20 was cloned from *E. coli* MRE600 genomic DNA into pET24b vector (Novagen).[20] Site-directed mutagenesis was used to introduce cysteine residues at three non-conserved positions (Ser 13, Ser 23, and Lys 49). The mutation was confirmed by sequence analysis. The mutated S20 proteins were expressed individually in *E.coli* pRBL21 and purified as described for wild-type protein.[20] The fourth modified S20 (Ile 55) was available.

www.manaraa.com

Derivatization of S20 proteins. The fluorescent reagent 7-diethylamino-3-((4′-(iodoacetyl)amino)phenyl)-4-methylcoumarin (DCIA; Molecular Probes) was used to test the accessibility of each introduced cysteine residue for derivatization. Non-specific derivatization at positions other than the cysteine residues was assessed by using the wild-type S20, in parallel with the cysteine-containing S20 mutant proteins. Derivatization of cysteine-containing S20 proteins by DCIA and Fe(II)–BABE(1-(p-bromoacetamidobenzyl) ethylenediaminetetraacetate) was done as described.[21]

Formation of Fe(II)–S20 containing RNPs. The natural 16S rRNA, isolated from natural 30S subunits as described previously,[35] was pre-incubated at 42°C for 15 minutes in buffer A (20 mM K^+-Hepes (pH 7.6), 20 mM $MgCl_2$). The Fe(II)–S20 containing RNPs were formed by mixing the 16S rRNA (40 pmoles) with the Fe(II)-S20 (200 pmoles), and the other necessary r-proteins (240 pmoles each). The KCl concentration was adjusted to 330 mM in each of the reactions in a final volume of 100 µl, and they were incubated at 42°C for 60 minutes. The reaction mixtures were kept on ice for 10 minutes before purification on spin columns and directed probing (see below).

Purification of RNPs from Fe(II)–S20. The complexes containing Fe(II)–S20 proteins were purified to remove any unbound protein by using spin columns, prior to directed hydroxyl radical probing. This was done as described by Culver and Noller,[21] the only difference being the centrifugation speed, which in our case was 6. The purified RNPs were kept on ice for 10 minutes before the hydroxyl radical probing.

The 30S subunits were also purified using centrifugal filter devices Microcon YM-50, at a speed of 6.4 rpm. The reaction mixture volume was reduced to half (50 µL), followed by three washes with 400 µL of buffer A.

www.manaraa.com

Directed hydroxyl radical probing. Directed hydroxyl radical probing of 16S rRNA from Fe(II)–S20, in the different RNPs and the subsequent primer extension analysis was done as described by Culver and Noller.[21]

Results

Expresion, purification and derivatization of cysteine containing S20. The cysteine containing S20 proteins were expressed and purified as described,[33] and the availability of the cysteine residue for derivatization was assessed using the fluorescent label 7-diethylamino-3-((4′-(iodoacetyl)amino)phenyl)-4-methylcoumarin[21] (DCIA) (data not shown). Derivatization with Fe(II)-BABE (1-(p-bromoacetamidobenzyl) ethylenediaminetetraacetate) and purification of the r-protein was performed using published methods.[21] The dervatization with Fe(II)-BABE was also confirmed and assessed by the fluorescence assay (data not shown).

Purification and probing of RNPs containing Fe(II) derivatized S20 protein. RNPs containing the derivatized S20 protein are formed, and purified by size exclusion chromatography. It is very important to remove any free modified protein from the reaction mixture, since its presence might result in production of spurious hydroxyl radicals, and subsequent nonspecific cleavage. Traditionally, 30S subunits are purified from unbound material by ultracentrifugation through sucrose gradients. The 30S subunits used for probing were purified through sucrose gradients or with centrifugal filter device Microcon YM-50. Base-specific footprinting with DMS shows that the 16S rRNA in 30S subunits containing Fe(II) tethered S20 purified using the centrifugal filter device Microcon YM-50, are folded similarly to 16S rRNA in the unpurified 30S subunits (data not shown).

www.manaraa.com

After purification, hydroxyl radicals are generated. Hydroxyl radicals cleave the RNA backbone proximal to the Fe(II)-modified sites and cleavage sites are identified by primer extension. The results are mapped on the secondary structure of 16S rRNA and on the tertiary structure of 16S rRNA from the 30S subunit of the *E. coli* ribosome.

Directed hydroxyl radical probing of the minimal complexes Fe(II)–S20 /16S rRNA. The four minimal complexes containing Fe(II) derivatized S20 r-proteins have different cleavage patterns (Figure 3, Table 1), but the cleavage sites for each one of them are localized in the 5' domain and helix 44 of the 16S rRNA (Figure 4b-e).

In the RNP containing C13S20, cleavage sites from the hydroxyl radicals are present in helices 5, 6, 8, 13, 14 (5' domain) and also at two discrete sites in helix 44 (Figure 4b). The most extensive cleavage by the radicals generated from Fe(II)-C13S20, in the minimal complex, is observed in helix 8 and at the junction formed by helices 5, 6 and 14.

For the minimal complex containing C22S20, the cleavage is the most pronounced in the 5' domain, when compared to the other minimal RNPs (Figure 3a, Figure 4). Helices 6, 8, 9, 11, 13 and 14, from the 5' domain, all show cleavage in different degrees and weak cleavage, at two different sites is also observed in helix 44 (Figure 4c). In the minimal complex containing Fe(II)C22S20, helices 8 and 13 showed the most extensive cleavage.

When C47S20 is used as a probe in the minimal complex, the cleavage pattern was quite different from all the other three derivatized proteins (Figure 3, Figure 4), which is consistent with the positioning of residue 47 in the r-protein S20 (Figure 2e, f). In the minimal complex C47S20/16S rRNA, helices 8, 9, 11, and loop 360 are cleaved in

www.manaraa.com

the 5' domain, while in helix 44 there are three discrete cleavage sites (Figure 4d). Some nucleotides around 1440 are weakly targeted, while stronger cleavage takes place in the regions 1450 and 1460 (Figure 3c, Figure 4d). The cleavage observed in helix 44 from the hydroxyl radicals generated by Fe(II) tethered C47S20 is the strongest from all the minimal complexes (Figure 3d).

The cleavage pattern observed in the complex containing C55S20 as a probe is very similar to the one observed in the C22S20/16S rRNA RNP, but with lower intensity (Figure 3, Figure 4c, e). As it was mentioned earlier, residues 22 and 55 are located on two different helices of S20, but they are very close in space (Figure 2e, f). Especially in helices 8 and 13 the weaker cleavage is obvious, and some of the nucleotides cleaved from Fe(II)-C22S20 are not cleaved from C55S20 (Figure 4c,e). Very weak cleavage form Fe(II)-C55S20 is observed in helix 44 (Figure 3c, Figure 4e), and probably that is justified since on the tridimensional structure of 16S rRNA with S20, residue 55 seems to be more buried than residue 22.

Directed hydroxyl radical probing of 30S containing Fe(II)–S20. The cleavage patterns for each of the 30S subunits containing one of the Fe(II)-S20 protein are different, but when compared to the corresponding minimal complexes, it is obvious that they are very similar (Table 1, Figure 3, Figure 5). Also the cleavage patterns observed in the 30S subunits when S20 is used as a probe are very similar to the ones previously published, with a few exceptions: the extent and intensity of cleavage are higher in the experiments performed in this study, which may arise from a better quality of the starting materials, like commercially available BABE or fresh hydrogen peroxide; in one of the gels previously published[33] two of the samples were swapped (C13S20-30S and C22S20-

www.manaraa.com

30S in the primer extension with primer 232); and it seems that in the original experiments the extent of derivatization of C47S20 was lower than the one obtained in this study, which led to less cleavage, which again can be due to the quality of BABE used. At the time when probing was performed in the 30S subunit using r-protein S20 as a probe, the fluorescence assay was not yet used to assess the availability of cysteine or the derivatization of S20.[33] These conclusions arise after extensive control experiments, and they were also consistent with the mapping of the data on the three dimensional structure of 16S rRNA from the 30S subunit, which is now available (Figure 5).

The majority of differences observed between the cleavage patterns in the 30S subunits and the corresponding minimal complexes are usually in intensity, and not in the position of the cleavage sites. In a few regions it seems that there are some subtle differences, like new cleavage sites appear or even more interestingly, disappear (Table 1, Figure 3, Figure 5). The question that comes up is are these subtle differences real or they are inherent small differences observed because the probing of distinct particles was performed at different times.

Directed hydroxyl probing of RNPs of different complexities containing Fe(II)-C13S20. To address this question, a different approach was taken. RNPs of increasing complexity were probed and the same derivatized S20 was used as a probe in each one of them. Small differences and similarities are made more credible by this approach. The results are presented for two of the derivatized S20 protein.

To have up most confidence in subtle differences, in the experiments in which RNPs of difference complexities are probed with Fe(II)-S20 protein, all the RNPs, including 30S subunits, are purified by size exclusion. Base-specific footprinting with

www.manaraa.com

DMS shows that the 30S subunits containing Fe(II) tethered S20 purified by size-exclusion, are the same as the non-purified 30S subunits (data not shown).

Six different RNPs of different composition are formed and probed with the same derivatized S20, in the same experiment for a better comparison. The first complex is S20/16S rRNA, which is mainly used as a control, since it lacks cysteine, Fe(II) should not be present and hydroxyl radicals should not be generated. The other complexes are Fe(II)-C13S20/16S rRNA, Fe(II)-C13S20/1^{o}/16S rRNA (1^{o} mixture contains S4, S7, S8, S15 and S17), Fe(II)-C13S20/1^{o}/S12+S16+S5/16S rRNA, Fe(II)-C13S20/5'/16S rRNA (5' contains S4, S8, S17, S16, S12, S5) and Fe(II)-C13S20-30S. Since this was an exploratory experiment, more complex particles were chosen to be able to see changes. The minimal RNP is the starting point for the comparison; it is the least complex particle that can be formed. The next RNP contains all the primary r-proteins that interact directly with 16S rRNA and start organizing their corresponding domains, indicated in the appropriate color on the assembly map[16, 17] (Figure 1b, Figure 6). The primary r-proteins S4 and S17 bind directly to the 5' domain (shown in red in Figure 1), S8 which binds to the central domain (shown in green in Figure 1) will promote the binding of a secondary binding protein that binds to 5' domain, S15 binds to the central domain (red) and S7 initiates the assembly of the 3' major domain[36] (shown in blue) (Figure 1). The RNP containing all the primary r-proteins and the three secondary binding proteins will show the importance of the secondary binding r-proteins that bind in the 5' domain (shown in red in Figure 1b), by comparison with the RNP containing just the primary binding r-proteins. The RNP containing the 5' domain proteins (Figure 1b, Figure 6) will show if organizing the 5' domain is enough to obtain a cleavage pattern similar to the one in the

www.manaraa.com

30S subunit. Also, it will allow inferring if S15 and S7 are important for the organization of 16S rRNA surrounding S20, by contrast with the previous particles. Finally, the 30S subunit is the product of the assembly and another important evaluation point. This approach allows direct comparison of the cleavage patterns from the same probe in different RNPs and it is possible to deconvolute which proteins produce changes in the environment of the probe.

The cleavage patterns are quite similar for all the RNPs containing C13S20 mentioned above (Figure 7), with a few notable exceptions. Most importantly, the cleavage intensity is lower for the minimal complex containing Fe(II)-C13S20 (Figure 7) than in all the other RNPs, especially in the 5' domain of 16S rRNA. There are a few cleavage sites in which the change in intensity is more pronounced as more proteins are added, than others but we will not focus on the cleavage sites that become more intense, since that is a general phenomenon for this probe, but on three sites where the intensity varies between the RNPs. R-protein dependent variation allow determination of roles for specific r-proteins, and implicitly of r-proteins that are important in folding distinct regions of 16S rRNA.

The most intriguing cleavage region from Fe(II)-C13S20 is the one around nucleotide 80 (Figure 6a), in helix 6 of 16S rRNA where for the RNP containing all primary proteins and the one containing all the primary and the secondary S16, S5 and S12 r-proteins the cleavage becomes more pronounced than in the minimal complex, but for the RNP containing the 5' domain mix there is no increase in the cleavage intensity compared to the minimal complex (Figure 6a). In the 30S subunit the cleavage is the

www.manaraa.com

most intense, from all the RNPs probed. From these experiments it can be concluded that the change is brought by the presence of the primary binding r-proteins S7 and S15.

Interestingly, there are two cleavage sites where the intensity initially increases as all the other primary r-proteins are added and decreases in the more complex RNPs (Figure 6a, c) and these are the regions around nucleotides 60 and 350. While dispersed in the primary sequence, in the secondary structure of the 16S rRNA these nucleotides are located very close to each other, at a helical junction formed by helices 5, 6, 13 and 14. Thus they are probably influenced by assembly of the same r-protein(s).

Directed hydroxyl probing of RNPs of different complexities containing Fe(II)-C22S20. The same particles are also investigated using C22S20 as a probe. In this case, the cleavage patterns and intensities obtained for the different RNPs are very similar in the 5' domain (Figure 8a-c), with one exception. Interestingly, the exception is the cleavage region close to nucleotide 80 (Figure 8a), in helix 6, and the same behavior was observed when C13S20 is used as a probe (see above). A stronger cleavage than in the minimal particle is observed at the addition of the primary proteins or the primary and S16, S5 and S12, but no increase when only the 5' domain proteins are added, while in the 30S subunit the cleavage is the most pronounced. Much more intriguing are the cleavage sites observed in helix 44 from Fe(II)-C22S20 (Figure 8d). Three discrete cleavage sites are observed in the minimal complex and the 30S subunit in regions 1430, 1450 and 1460. The intensity is quite lower in the minimal complex compared to the fully assembled 30S subunit, and the particles of intermediate complexity show varied intensities. These results are suggesting that helix 44 interacts differently with S20 during

www.manaraa.com

the assembly, compared to the interaction from the minimal complex and the 30S subunit.

Discussion

The cleavage patterns observed for the minimal RNPs containing Fe(II)-tethered to S20 are very similar to the ones observed in the fully assembled 30S subunit. The placement of the four different residues on S20 makes possible a thorough examination of the surroundings of this r-protein. Probing with hydroxyl radicals generated from Fe(II)-tethered through residues 13 and 47 are very different from each other and the other derivatized S20 proteins. The cleavage from the S20 derivatized at positions 22 and 55 are quite similar (Figure 4c, e), as expected from their relative placement in the r-protein S20(Figure 2e, f). The results of the directed hydroxyl probing in the minimal complexes correlate well with the relative positions of the derivatized residues in S20 and show that each minimal complex has its own cleavage pattern (Figure 4). The representation of the cleavage sites on the 16S rRNA from the *E. coli* 30S subunit[1] illustrates how much information can be obtained by directed hydroxyl radical footprinting, and confirms the validity of the experimental data (Figure 5 a, b, e, f, i, j, m, n). The cleavage sites and their intensity correlate very well with the placement of the respective probe in the structure of S20 and of the 30S subunit (Figure 5 a, b, e, f, i, j, m, n). For each of the Fe(II)-S20 protein the intensity of cleavage is highest around the Fe(II) ion and it decreases as the distance increases, due to the limited range of the hydroxyl radicals. Even in then minimal complexes containing Fe(II) derivatized S20, helix 44 is positioned very similarly to its position in the 30S subunit.

www.manaraa.com

The comparison with the cleavage pattern in the corresponding 30S subunits demonstrate that both in the minimal complex and the final product of the assembly, the surroundings of the probe r-protein S20 are fairly similar (Figure 5). As it was shown by base-specific footprinting and solution hydroxyl radical probing S20 interacts with both the 5' domain and helix 44 of the 3' minor domain of 16S rRNA in both the minimal complex and the fully assembled 30S subunit. Beside that, it was also shown that even in the absence of any r-proteins the 5' domain has an architecture that resembles closely the architecture of the 30S subunit.[37]

Fe(II)-C13S20 as a probe for the assembly of the 30S subunit - general trends. The cleavage patterns in the minimal complex and the 30S subunit when C13S20 is used as a probe are similar, yet distinct. This suggests that even in the presence of only r-protein S20 both the 5' domain and helix 44 are sampling 30S-like positions relative to S20 in the C13S20/16S rRNA complex. At the assembly of other r-proteins adjustments are taking place in the conformation of 16S rRNA surrounding S20 protein. Especially notable is that in the minimal complex the intensity of cleavage is lower compared to all other RNPs studied, at all cleavage sites, with one exception (region 80).

Residue 13 is located in the N-terminal helix of S20, the first and longest helix from the three that compose S20 (Figure 2e and f). The two other helices that are present in S20 are of very similar size, while the N-terminus helix is longer. Residue 13 is located in a region of the N-terminal helix that is not stacked with the other helices. This region can probably move more than other regions of S20, and it seems that in the minimal complex is not as close to the 16S rRNA as in other more complex RNPs or it is more dynamic, as the lower intensity of the cleavage is showing. Probably a combination of the

www.manaraa.com

dynamics of the N-terminal helix of the r-protein S20 and the 16S rRNA surrounding it, account for the lower intensity of cleavage in the minimal complex. When more r-proteins are added (all the other primary proteins for example) the 16S rRNA around S20 adopts an architecture more similar to the one in the 30S subunit. The decrease in intensity could be again attributed to the protection of 16S rRNA by r-proteins, but all of the r-proteins are quite far away from S20.

The differences observed between the different RNPs explored make it possible to identify some of the r-proteins responsible for the changes of 16S rRNA conformation around r-protein S20. Interestingly, the differences between the minimal complex and the other complexes are more intense in the 5' domain than in the 3' minor domain (helix 44).

Role of the primary r-proteins in organizing the 16S rRNA surrounding of Fe(II)-C13S20. The addition of all the other primary r-proteins organizes the 16S rRNA surrounding the C13S20 probe and makes the cleavage pattern more similar to the one in the 30S subunit than the one observed in the minimal complex (Figure 7a-c). From our results it can be inferred that probably even only in the presence of the other primary r-proteins that bind the 5' domain (S4 and S17), the cleavage pattern is going to be very similar to the one in the fully assembled 30S subunit.

Interestingly, even the presence of r-proteins S7 and S15 makes a difference in the organization of the surroundings of S20, though they bind different domains of 16S rRNA. In the structure of the 30S subunit, S20 is located at the bottom of the body, S7 is located in the head and S15 is located somewhere in the middle, in the platform (Figure 6). The intensity of cleavage in the region of nucleotide 80 (Figure 7a) is similar to the

www.manaraa.com

one in the 30S subunit if the RNP contains the r-proteins S7 and S15. Only the presence of the primary and secondary r-proteins that bind the 5' domain is not sufficient to acquire an architecture of 16S rRNA similar to the one in the 30S subunit. It was shown that the domains can assemble independently of each other, in the presence of the required r-proteins,[38,39] but obviously there are still interdomain adjustments that take place and require the presence of r-proteins from other domains.

Role of the secondary protein S16 in organizing the surroundings of Fe(II)-C13S20. The addition of the secondary proteins S16, S5 and S12 also brings a few changes (see regions 60 and 350, Figure 7a, c), that augment the organization of the 16S rRNA surrounding the r-protein S20. The binding of all the primary proteins to the 16S rRNA increases the intensity of cleavage at the multiple-stem junction of helices 5, 6, 13 and 14 (Figure 9a) but as more proteins are added, the cleavage intensity decreases. For the same junction, in the presence of all the primary and secondary proteins that bind to the 5' domain, but in the absence of the primary r-proteins S7 and S15, the intensity of cleavage is the same as in the 30S subunit. By elimination, the proteins that can be responsible for the decrease in intensity are S5, S12 and S16. When analyzing the cleavage sites mentioned above (regions 60 and 350) on the three dimensional structure of 16S rRNA from the 30S subunit[1] and the three proteins that might be responsible for the effect it is obvious that S16 is the most likely effector (see Figure 9b). The influence of S16 on the cleavage pattern can be explained in two ways, since the intensity of cleavage decreases. S16 is protecting the 16S rRNA from the hydroxyl radicals generated from C13S20, or the rRNA is rearranged at the binding of S16 and the distance between the probe and the rRNA is increasing. From the positioning of S16 in the structure of the

www.manaraa.com

30S subunit (Figure 9b), the most plausible is that S16 is going to protect the rRNA from the hydroxyl radicals. Probably in the RNP containing Fe(II)-C13S20, the other primary r-proteins, and the secondary r-proteins S5, S12 and S16, the 16S rRNA has a structure which is the most similar to the one in the 30S subunit, around the r-protein S20.

Fe(II)-C22S20 as a probe for the assembly of the 30S subunit. The results obtained from probing different complexes using as a probe Fe(II)-C22S20 are quite different from the ones already discussed for Fe(II)-C13S20. The comparison of the cleavage patterns in different RNPs containing Fe(II)-C22S20 reveals that the differences in the 5' domain are minimal between particles (Figure 8a-c), while they are quite significant in the 3' minor domain (Figure 8d).

Cleavage in the 5' domain from Fe(II)-C22S20. The different RNPs that contain Fe(II)-C22S20 as a probe show very similar patterns of cleavage in the 5' domain (Figure 8a-c). The only exception is in helix 6, region 80, where some notable differences are observed in the intensity of cleavage between the different RNPs (Figure 8a). Remarkably, the same behavior was observed for the particles containing Fe(II)-C13S20, and it was discussed in the previous section (see above). It seems that the 16S rRNA close to the derivatized cysteine residue 22 has almost the same conformation in the 5' domain in the minimal complex and all the other RNPs, the final product of assembly, the 30S subunit, included.

Cleavage in helix 44 from Fe(II)-C22S20. The comparison of cleavage patterns for Fe(II)-C22S20 containing RNPs in helix 44 shows quite a few differences (Figure 8d). Interestingly, the minimal complex and fully assembled 30S subunit have the most similar cleavage pattern, although for the minimal complex the intensity is lower.

www.manaraa.com

Probably 16S rRNA adopts very similar conformations around the probe in both the minimal complex and the 30S subunit. All the other RNPs examined, with intermediate composition, show almost no cleavage. Thus, the complexes with intermediate composition are quite different from the Fe(II)C22S20/16S rRNA complex and the final product of assembly, the fully formed 30S subunit. The two possible causes for the observed differences are changes in dynamics or quenching. It is probable that in the complexes that contain more r-proteins helix 44 is more dynamic and cleavage does not take place significantly. Or, quenching can take place from another protein or an element of the 16S rRNA, but it is a transient state, since it is not present in the 30S subunit. Thus, the most plausible explanation for our results is the increase in the dynamics, but further experiments will hopefully bring a better understanding.

One of our hypotheses was that we will be able to observe differential interaction of r-protein S20 with its two binding domains during assembly of the 30S subunit with our approach. Even though the interaction of S20 with the 5' domain is almost the same in all of the RNPs explored, significant differences were observed for the 3' minor domain, suggesting that indeed along the assembly S20 interacts differentially with its two binding domains, 5' domain and helix 44.

Conclusions and future prospects

R-protein S20 can be used as a probe to study the assembly of the 30S subunit, as it is shown by our results. The study of RNPs of different complexities separately can give information on the assembly, but it is much easier and informative to probe and analyze RNPs of different compositions in the same time. The two probes that were used to explore RNPs of different compositions gave different types of information and

www.manaraa.com

showed the versatility of our approach. When Fe(II)-C13S20 is used as a probe in different RNPs, differences which are not obvious at the comparison of the minimal complex and 30S subunit separately emerge, along with roles for some r-proteins in the organization of the 16S rRNA around S20. For Fe(II)-C22S20 differential interaction with 5' domain and helix 44 during the assembly are observed. Other RNPs containing Fe(II)-C13S20 and Fe(II)-C22S20 will also be explored since some differences observed in the complexes mentioned in this study cannot be attributed to only one protein, though inferences can be made. The study of different RNPs containing Fe(II)-C47S20 or Fe(II)-C55S20 will complete the picture of the assembly of the 30S subunit from the point of view of r-protein S20.

References

1. Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H., *Science* **2005,** 310, (5749), 827-34.
2. Yusupov, M. M.; Yusupova, G. Z.; Baucom, A.; Lieberman, K.; Earnest, T. N.; Cate, J. H.; Noller, H. F., *Science* **2001,** 292, (5518), 883-96.
3. Brodersen, D. E.; Clemons, W. M., Jr.; Carter, A. P.; Wimberly, B. T.; Ramakrishnan, V., *Acta Crystallogr. D Biol. Crystallogr.* **2003,** 59, (Pt 11), 2044-50.
4. Brodersen, D. E.; Clemons, W. M., Jr.; Carter, A. P.; Wimberly, B. T.; Ramakrishnan, V., *J. Mol. Biol.* **2002,** 316, (3), 725-68.
5. Wimberly, B. T.; Brodersen, D. E.; Clemons, W. M., Jr.; Morgan-Warren, R. J.; Carter, A. P.; Vonrhein, C.; Hartsch, T.; Ramakrishnan, V., *Nature* **2000,** 407, (6802), 327-39.

www.manaraa.com

6. Agalarov, S. C.; Sridhar Prasad, G.; Funke, P. M.; Stout, C. D.; Williamson, J. R., *Science* **2000,** 288, (5463), 107-13.

7. Wimberly, B. T.; White, S. W.; Ramakrishnan, V., *Structure* **1997,** 5, (9), 1187-98.

8. Golden, B. L.; Hoffman, D. W.; Ramakrishnan, V.; White, S. W., *Biochemistry* **1993,** 32, (47), 12812-20.

9. Clemons, W. M., Jr.; Davies, C.; White, S. W.; Ramakrishnan, V., *Structure* **1998,** 6, (4), 429-438.

10. Davies, C.; Gerstner, R. B.; Draper, D. E.; Ramakrishnan, V.; White, S. W., *EMBO J.* **1998,** 17, (16), 4545-4558.

11. Noller, H. F.; Green, R.; Heilek, G.; Hoffarth, V.; Huttenhofer, A.; Joseph, S.; Lee, I.; Lieberman, K.; Mankin, A.; Merryman, C.; et al., *Biochem. Cell. Biol.* **1995,** 73, (11-12), 997-1009.

12. Traub, P.; Nomura, M., *J. Mol. Biol.* **1968,** 34, (3), 575-93.

13. Noller, H. F.; Woese, C. R., *Science* **1981,** 212, (4493), 403-11.

14. Cannone, J. J.; Subramanian, S.; Schnare, M. N.; Collett, J. R.; D'Souza, L. M.; Du, Y.; Feng, B.; Lin, N.; Madabusi, L. V.; Muller, K. M.; Pande, N.; Shang, Z.; Yu, N.; Gutell, R. R., *BMC Bioinf.* **2002,** 3, 2.

15. Mizushima, S.; Nomura, M., *Nature* **1970,** 226, (5252), 1214-18.

16. Grondek, J. F.; Culver, G. M., *RNA* **2004,** 10, (12), 1861-1866.

17. Traub, P.; Nomura, M., *Cold Spring Harb. Symp. Quant. Biol.* **1969,** 34, 63-7.

18. Held, W. A.; Ballou, B.; Mizushima, S.; Nomura, M., *J. Biol. Chem.* **1974,** 249, (10), 3103-11.

www.manaraa.com

19. Traub, P.; Nomura, M., *Proc. Natl. Acad. Sci. U.S.A.* **1968,** 59, (3), 777-84.

20. Culver, G. M.; Noller, H. F., *RNA* **1999,** 5, (6), 832-43.

21. Culver, G. M.; Noller, H. F., *Methods Enzymol.* **2000,** 318, 461-75.

22. Newcomb, L. F.; Noller, H. F., *RNA* **1999,** 5, (7), 849-55.

23. Joseph, S.; Noller, H. F., *Methods Enzymol.* **2000,** 318, 175-90.

24. Culver, G. M.; Heilek, G. M.; Noller, H. F., *J. Mol. Biol.* **1999,** 286, (2), 355-64.

25. Lancaster, L.; Culver, G. M.; Yusupova, G. Z.; Cate, J. H.; Yusupov, M. M.; Noller, H. F., *RNA* **2000,** 6, (5), 717-29.

26. Joseph, S.; Whirl, M. L.; Kondo, D.; Noller, H. F.; Altman, R. B., *RNA* **2000,** 6, (2), 220-32.

27. Joseph, S.; Noller, H. F., *EMBO J.* **1996,** 15, (4), 910-6.

28. Jagannathan, I.; Culver, G. M., *J. Mol. Biol.* **2004,** 335, (5), 1173-1185.

29. Jagannathan, I.; Culver, G. M., *J. Mol. Biol.* **2003,** 330, (2), 373-383.

30. Powers, T.; Noller, H. F., *J. Biol. Chem.* **1995,** 270, (3), 1238-42.

31. Powers, T.; Daubresse, G.; Noller, H. F., *J. Mol. Biol.* **1993,** 232, (2), 362-74.

32. Culver, G. M., *Biopolymers* **2003,** 68, (2), 234-249.

33. Culver, G. M.; Noller, H. F., *RNA* **1998,** 4, (12), 1471-80.

34. Stern, S.; Changchien, L. M.; Craven, G. R.; Noller, H. F., *J. Mol. Biol.* **1988,** 200, (2), 291-9.

35. Moazed, D.; Stern, S.; Noller, H. F., *J. Mol. Biol.* **1986,** 187, (3), 399-416.

36. Nowotny, V.; Nierhaus, K. H., *Biochemistry* **1988,** 27, (18), 7051-5.

37. Adilakshmi, T.; Ramaswamy, P.; Woodson, S. A., *J. Mol. Biol.* **2005,** 351, (3), 508-19.

www.manaraa.com

38. Agalarov, S. C.; Zheleznyakova, E. N.; Selivanova, O. M.; Zheleznaya, L. A.; Matvienko, N. I.; Vasiliev, V. D.; Spirin, A. S., *Proc. Natl. Acad. Sci. U.S.A.* **1998,** 95, (3), 999-1003.

39. Weitzmann, C. J.; Cunningham, P. R.; Nurse, K.; Ofengand, J., *FASEB J.* **1993,** 7, (1), 177-80.

40. DeLano, W. L., *DeLano Scientific* **2002**, San Carlos, CA, USA.

www.manaraa.com

Table 1. Nucleotides at which cleavage by hydroxyl radicals is observed in the minimal complex and the 30S subunit, for each of the Fe(II)-derivatized S20 proteins.

Nucleotide	C13S20 /16S	C13S20 -30S	C22S20 /16S	C22S20 -30S	C47S20 /16S	C47S20 -30S	C55S20 /16S	C55S20 -30S
55 A	w	w						
56 U	w	w						
57 G	m	w						
58 C	m	w						
59 A	m	w						
60 A	w	w						
61 G	w	w						
99A		w						
100G	w	m						
101 A	m	m						
102 G	m	m	m	w				
103 U	w	w	m	w				
107 C	w	w						
108 G	m	m						
141 G					w	w		
142 G						w		
143 A					w	w		
144 G			w		w	m		
145 G			w		w	m	w	
146 G	w	w	m	w	w	m	m	w
147 G	w	w	s	m	w	m	s	m
148 G	w	w	v.s.	s	w	w	s	s
149 A	m	m	v.s.	v.s.			m	s
150 U	m	m	s	v.s			w	m
151 A	w	w	m	s				w
154 U			w					
155 A	w	w	m					
156 C	w	w	m					
157 U	w	w	m					
158 G	w	s	s	w				
159 G	w	s	s	w				
160 A	m	m	s					
161 A	w	w	w					
162 A	m	s	m					
163 C	v.s.	v.s.	s	m			w	w
164 G	v.s.	v.s.	s	s			w	w
165 G	s	v.s.	s	s			w	w
166 U	m	s	v.s.	s			m	m
167 A	w	m	s	m			w	w
168 G	w	m	s	m			w	w
169 C		w	m	w			w	w
170 U		w	v.s.	m			m	w
171 A		w	w	w			w	
174 A		w	m	w				
175 C	m	s	v.s.	s			m	w
176 C	s	v.s.	v.s.	v.s.	w		s	s
177 G	m	s	v.s.	v.s.	m	w	s	s

www.manaraa.com

Nucleotide	C13S20 /16S	C13S20 -30S	C22S20 /16S	C22S20 -30S	C47S20/ 16S	C47S20 -30S	C55S20 /16S	C55S20 -30S
178 C	m	s	s	v.s.	s	w	s	s
179 A	w	m	m	s	m	w	w	m
180 U		w	w	w	w		w	w
181 A			w		w			
182 A			w		w			
186 C					w	w		
187 G	w	w	w	w	m	w		w
188 C	w	w	w	w	s	m	w	w
189 A	w	w	m	w	s	w	w	w
190 A	w	w	m	w	v.s.	m	w	w
191 G	w	w	m	w	v.s.	s	m	m
192 A	w	w	m	w	v.s.	s	m	m
193 C	m	w	m	w	m	w	m	w
255 G					w	w		
256 U					w	w		
257 G					w	w		
258 G					w	w		
259 G								
260 G			m	m			w	w
261 U			w	w			w	w
267 C			m	m			w	w
268 U			m	m			w	w
269 C			m	m	w	w	w	w
270 A					w	w		
271 C					w	w		
272 C					w	w		
273 U								
274 A								
317 U			w					
318G			m	w				
319G			m	w				
320A			s	m				w
321A			s	s				w
322C			m	m				w
327A			w	w				
328C			w	w				
332G	w	w	m	w				
333U	w	w	m	w				
334C	s	w	s	s			m	m
335C	s	m	s	s			m	m
336A	m	m	m	m			w	w
337G	m	w	m	m			m	
344		w		w				
345C		w	w	w			w	
346G	w	m	s	m			w	w
347G	w	m	m	m			w	w
348G	m	m	w	w			m	w
349A	m	m	m	w			w	
350G	w	m	w	w				
351G	w	w						
352C	w							
353A	w							

www.manaraa.com

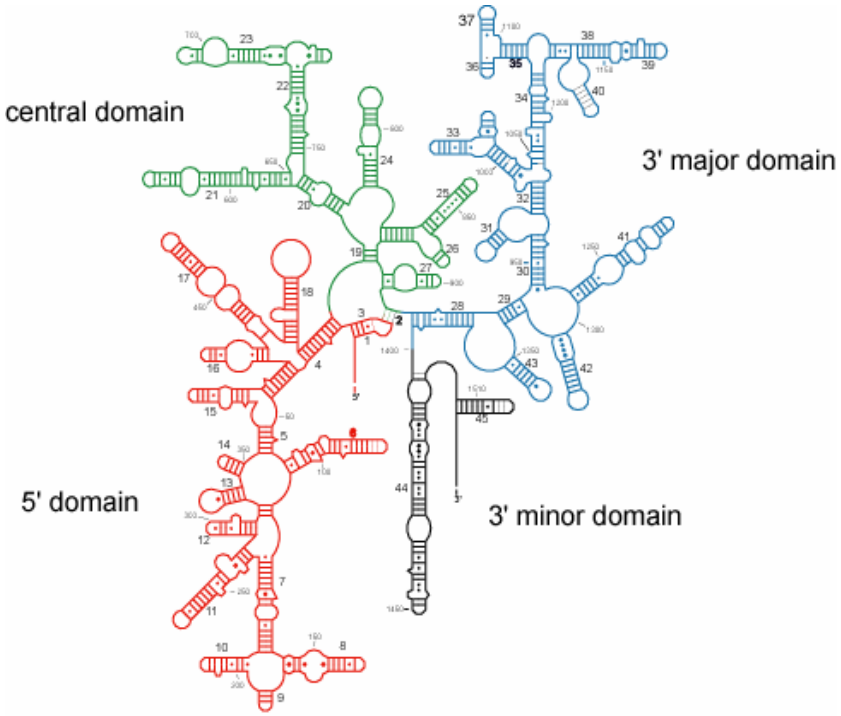
Nucleotide	C13S20 /16S	C13S20 -30S	C22S20 /16S	C22S20 -30S	C47S20/ 16S	C47S20 -30S	C55S20 /16S	C55S20 -30S
354G	w							
360G					w			m
361G					m			m
362G					m			m
363A					w			
960U	m							
961U	m							
962C	w							
1439G					w	w		
1440U					w	w		
1441A					w-m	w-m		
1444U	w	m	w	m				
1445U	m	m	w	m			w	w
1446A	w	w	w	m			w	w
1448C					m	w		
1449C					s	w		w
1450U					s	w	w	w
1451U					w		w	
1452C								
1453G					m	w		
1454G					m	w		
1455G					s	s		m
1456A					m	m		m
1457G	w	w			w			
1458G	m	m						
1459G	m	m	w	m				
1460C	m	s	w	s				
1461A		w		w				

w, m, s, v.s. indicate the intensity of the band at the specified nucleotide, and consequently the relative intensity of cleavage. w – weak, m – medium, s – strong, v.s. – very strong.

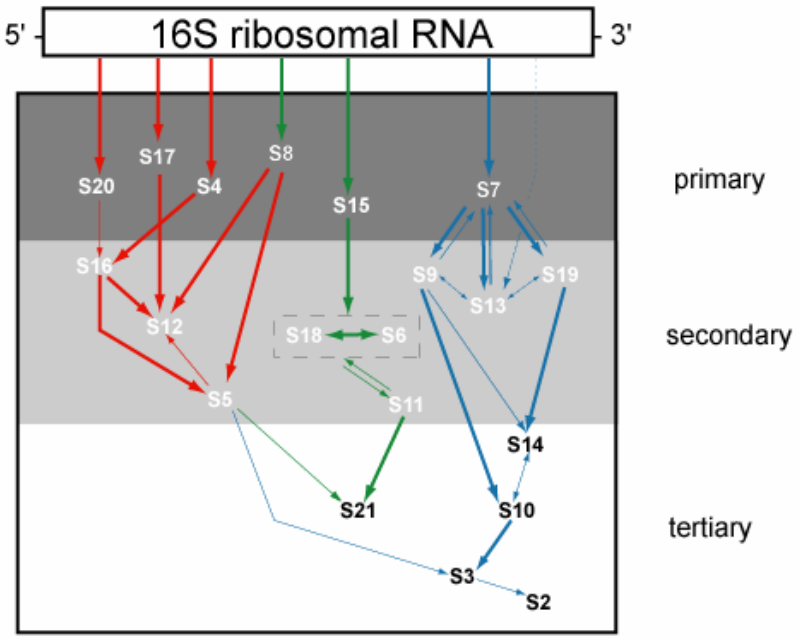
www.manaraa.com

Figure 1. 16S rRNA and the *in vitro* 30S subunit assembly map. (a) Secondary structure of 16S rRNA[14] with its domains in different color. In red is the 5' domain, in green the central domain, in blue the 3' major domain and in black the 3' minor domain. (b) in vitro assembly map of 30S subunit with the arrows connecting RNA and proteins, or proteins to proteins in the colors of the different domains to which they bind. The color coding of the arrows is the same as for (a). The r-proteins shown in white in the dark gray and the light gray regions are primary and respectively secondary binding r-proteins. The r-proteins shown in black on the white background are tertiary binding r-proteins, respectively. S6 and S18 are enclosed in a box to indicate that they bind as a heterodimer.

www.manaraa.com



(a)



(b)

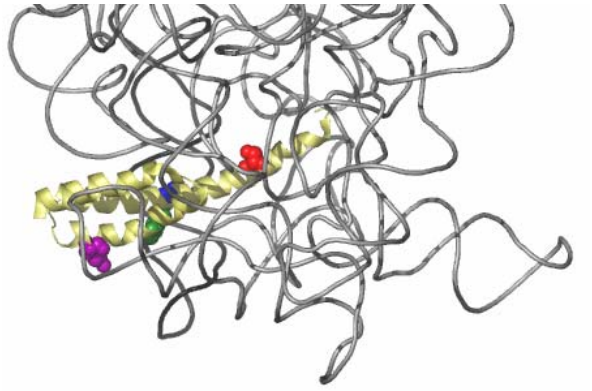
www.manaraa.com

Figure 2. Crystal structure of 16S rRNA and r-protein S20 from the *E. coli* 30S subunit.[1] (a) and (b) two views of the three dimensional structure of 16S rRNA and r-protein S20. 16S rRNA is shown in gray, with the 3' minor domain (helices 44 and 45) in black, and S20 in yellow. The cysteine substitution positions are shown: 13 in red, 22 in blue, 47 in pink and 55 in green. (c) and (d) Blow-up of the bottom portion of the 30S subunit. (e) and (f) two views of the structure of S20 (taken from (a) and (b) respectively) with sites of cysteine substitutions indicated. All figures containing 3-D structures were prepared using Pymol,[40] and the pdb file 2AW7.

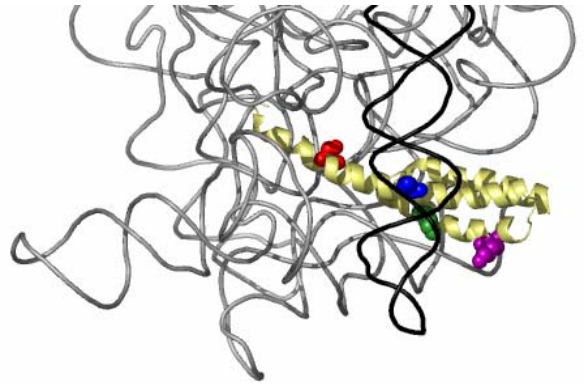
www.manaraa.com

(a)

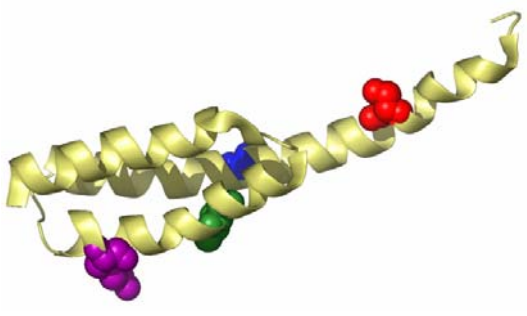
(d)



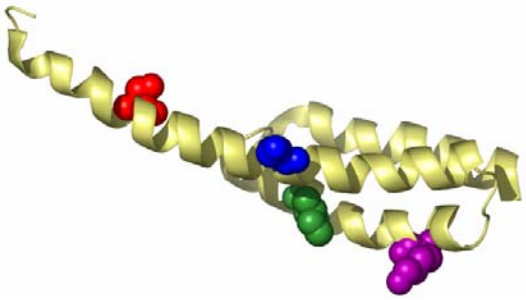
(b)



(e)



(c)

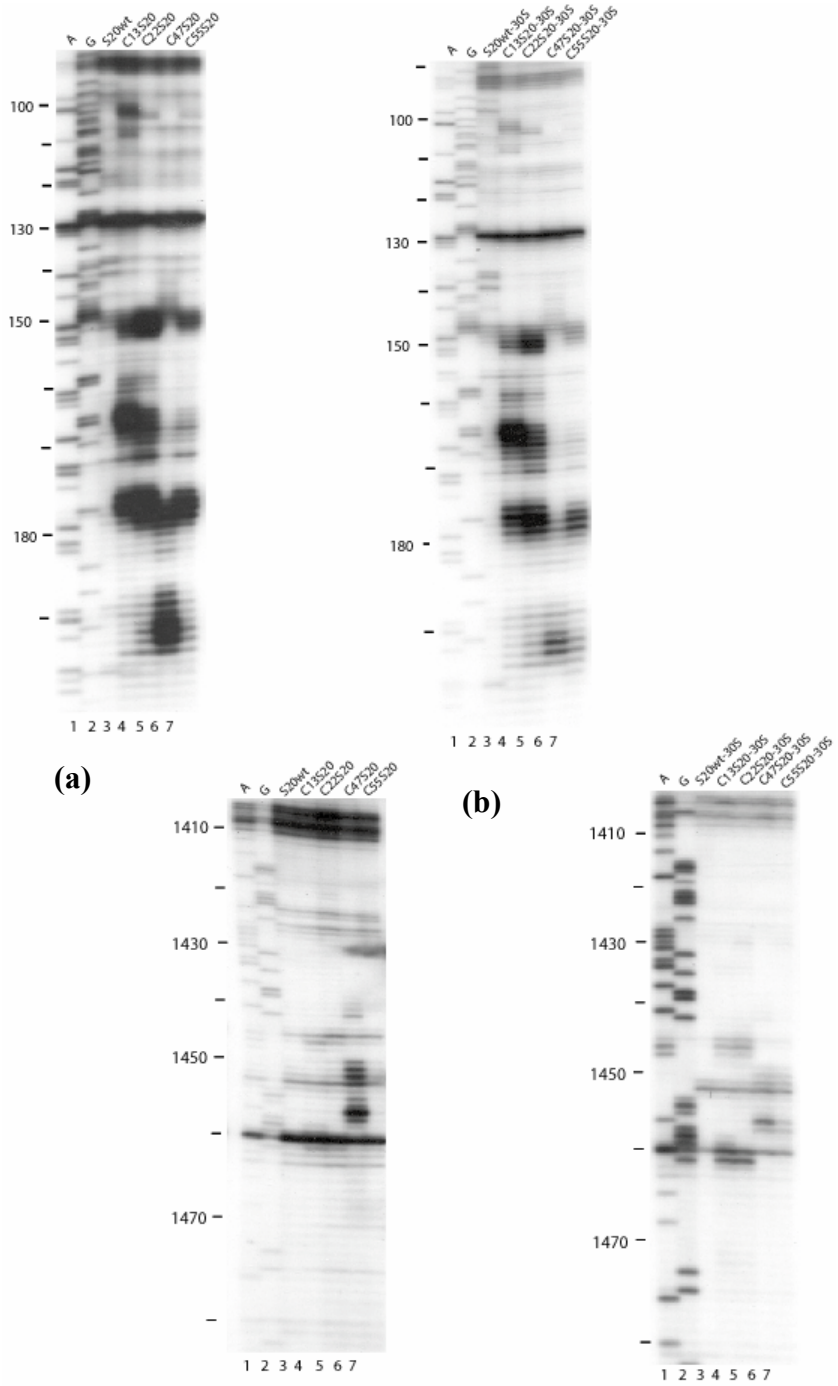


(f)

www.manaraa.com

Figure 3. Directed hydroxyl radical cleavage of 16S rRNA from Fe(II)–S20 in Fe(II)–S20/16S rRNA complexes and Fe(II)–S20-30S analyzed by primer extension. A and G are sequencing lanes. The other lanes are Fe(II)–S20/16S rRNA complexes (a and c) or Fe(II)–S20-30S (b and d) containing: wt S20 (lane 3), Fe(II)-C13S20 (lane 4), Fe(II)-C22S20 (lane 5), Fe(II)-C47S20 (lane 6), Fe(II)-C55S20 (lane 7). The primers used are 232 for (a) and (b) and 1508 for (c) and (d).

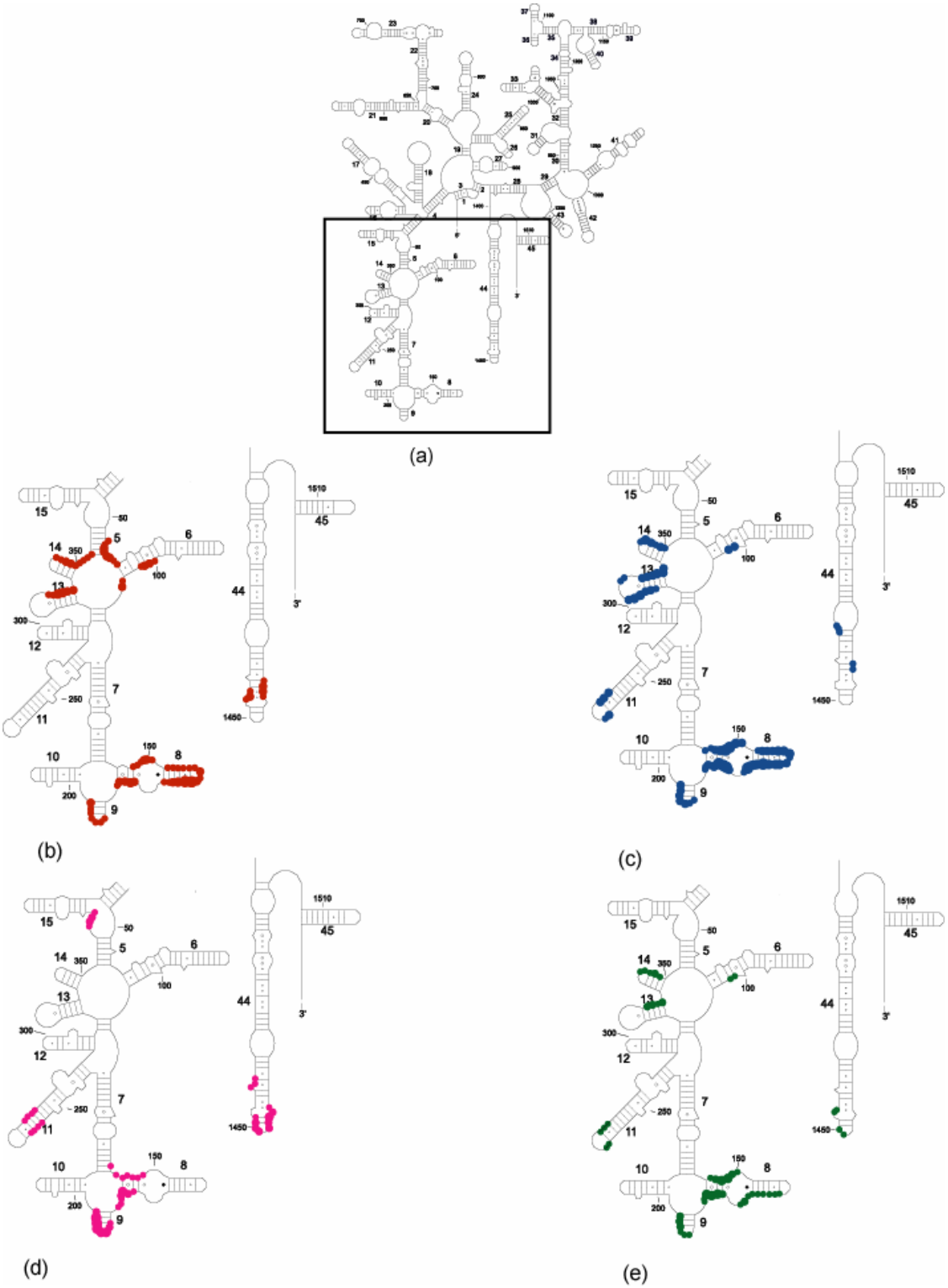
www.manaraa.com



www.manaraa.com

Figure 4. Hydroxyl radical cleavage sites shown on the secondary structure of 16S rRNA. (a) Secondary structure of 16S rRNA. The box contains the elements of the 5' domain and of the 3' minor domain used to represent cleavage sites in (b)-(e). The filled circles indicate positions of cleavage from Fe(II) derivatized S20 proteins in the minimal complex containing: (b) Fe(II)-C13S20, (c) Fe(II)-C22S20, (d) Fe(II)-C47S20 and (e) Fe(II)-C55S20. The color coding is the same as in Figure 2, C13S0 in red, C22S0 in blue, C47S20 in pink and C55S20 in green.

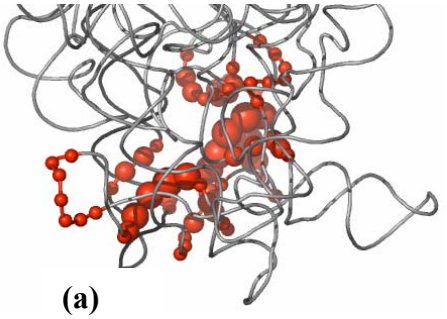
www.manaraa.com



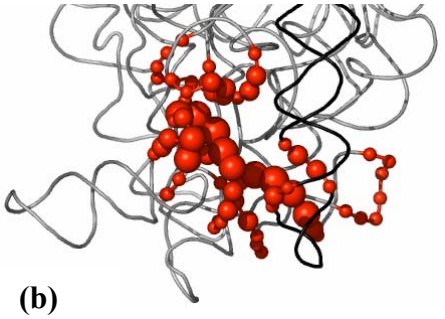
www.manaraa.com

Figure 5. Hydroxyl radical cleavage sites mapped on to lower part of the three-dimensional structure of 16S rRNA from *E. coli* 30S ribosomal subunit.16S rRNA is shown in gray, with the 3' minor domain (helices 44 and 45) in black, and S20 in yellow. The cleavage sites are represented as spheres, and the size of the sphere reflects the intensity of cleavage. For each Fe(II) derivatized S20 protein, there are four structures grouped for a better comparison: two views (180^{o} rotation) of the lower part of 16S rRNA with the cleavage sites from Fe(II) derivatized S20 protein for the minimal complex and two views (180^{o} rotation) for the fully assembled 30S subunit (the structure of the r-protein S20 with the specified residue in the appropriate color is included). (a)-(d) Fe(II)C13S20 (cleavage sites - red spheres): (a) and (b) minimal complex Fe(II)-C13S20/16S rRNA, (c) and (d) Fe(II)-C13S20-30S; (e)-(h) Fe(II)-C22S20 (cleavage sites - blue spheres): (e) and (f) minimal complex Fe(II)-C22S20/16S rRNA, (g) and (h) Fe(II)-C22S20-30S; (i)-(l) Fe(II)-C47S20 (cleavage sites - pink spheres): (i) and (j) minimal complex Fe(II)-C47S20/16S rRNA, (k) and (l) Fe(II)-C47S20-30S; and (m)-(p) Fe(II)-C55S20 (cleavage sites - green spheres): (m) and (n) minimal complex Fe(II)-C55S20/16S rRNA, (o) and (p) Fe(II)-C55S20-30S.

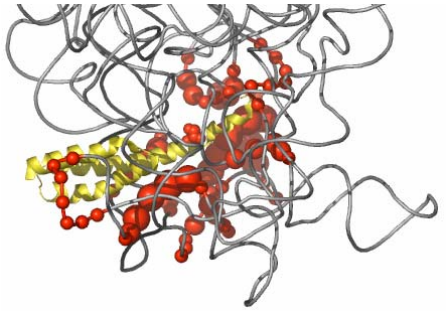
www.manaraa.com



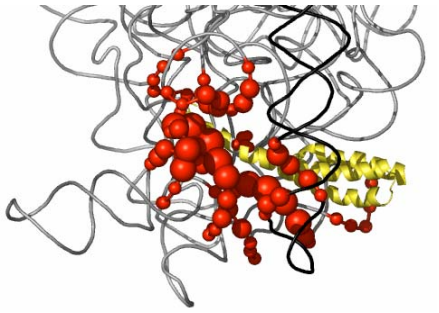
(a)



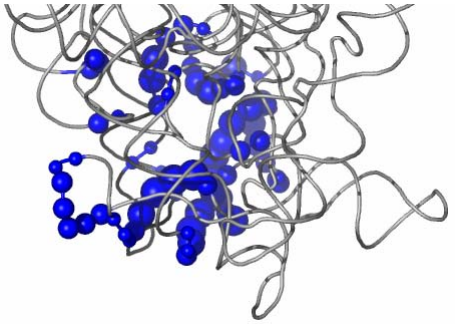
(b)



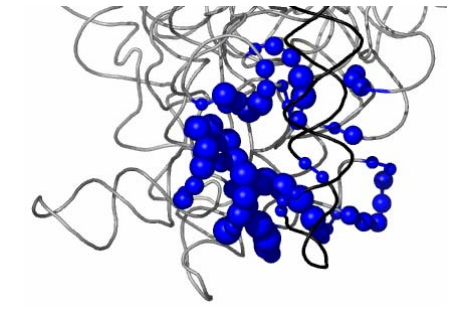
(c)



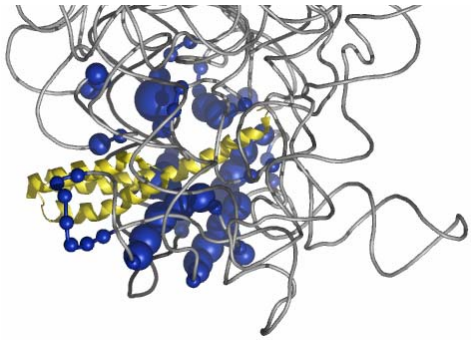
(d)



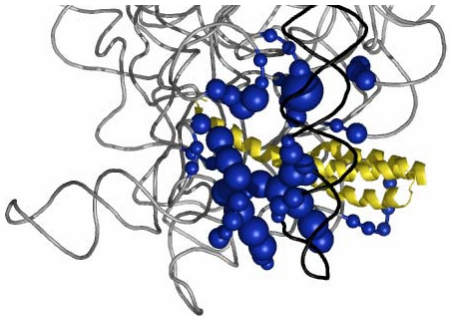
(e)



(f)

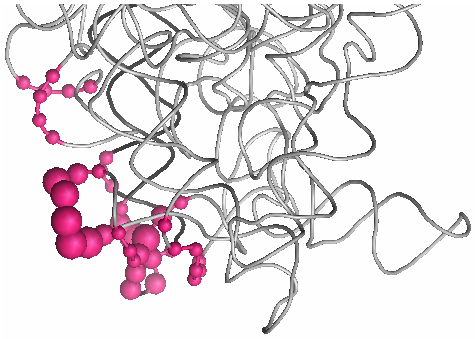


(g)

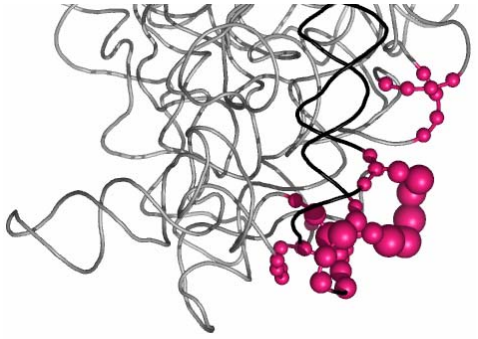


(h)

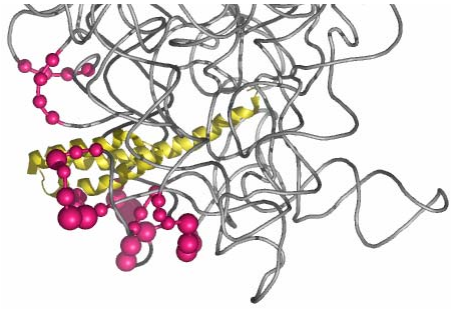
www.manaraa.com



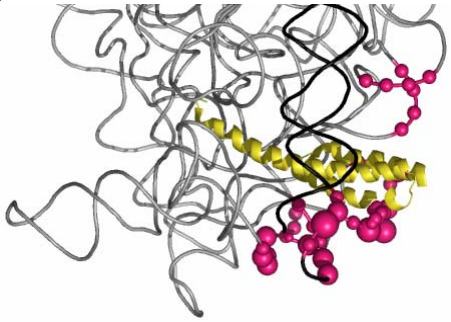
(i)



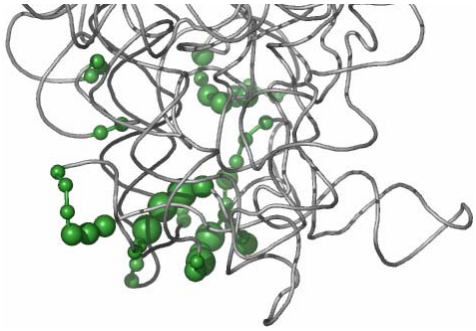
(j)



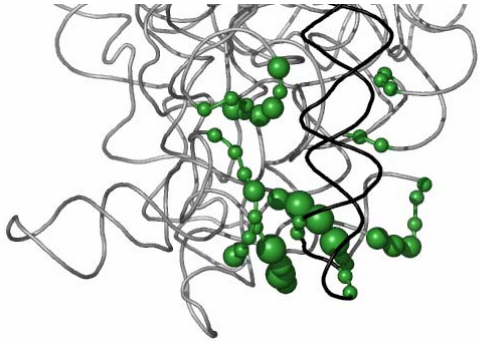
(k)



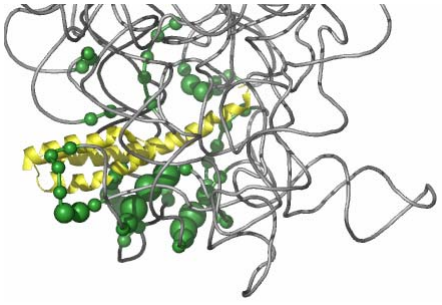
(l)



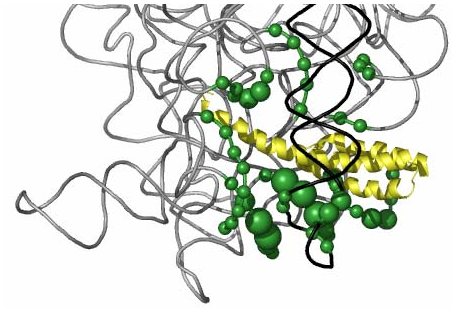
(m)



(n)



(o)



(p)

www.manaraa.com

Figure 6. Crystal structure of the 16S rRNA from the *E. coli* 30S subunit with all the primary r-proteins and the secondary binding r-proteins S5, S12 and S16. The 16S rRNA is shown in gray with the 3' minor domain (helices 44 and 45 in black), and the r-proteins are S4 green, S5 orange, S7 red, S8 pink, S12 brown, S15 lime green, S16 blue, S17 purple and S20 yellow. (a) solvent view of the 30S subunit. (b) 180^{o} rotation, interface view of the 30S subunit.

www.manaraa.com

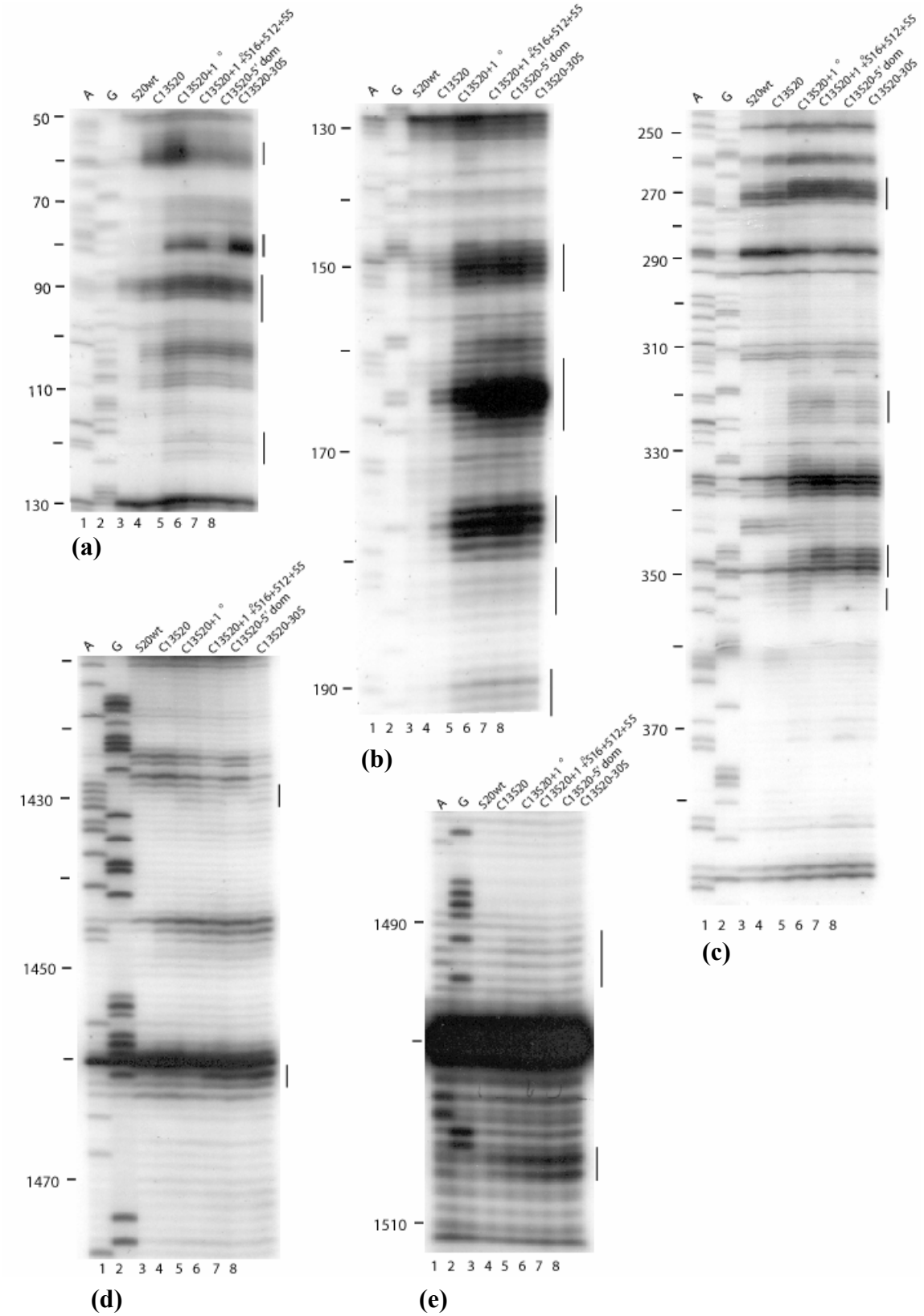
(a)

(b)

www.manaraa.com

Figure 7. Directed hydroxyl probing of RNPs of different complexities containing Fe(II)-C13S20 analyzed by primer extension. The sites of cleavage are indicated by the lines on the side of the gels. The lanes correspond to: sequencing lanes (lanes 1 and 2), and the complexes Fe(II)-C13S20/16S rRNA (lane 3), Fe(II)-C13S20/1^{o}/16S rRNA (1^{o} mixture contains S4, S7, S8, S15 and S17) (lane 4), Fe(II)-C13S20/1^{o}/S12+S16+S5/16S rRNA (lane 5), Fe(II)-C13S20/5'/16S rRNA (5'mixture contains S4, S8, S17, S16, S12, S5) (lane 6) and Fe(II)-C13S20-30S (lane 7). (a) primer 232 – upper part of the gel, (b) primer 232 – lower part of the gel, (c) primer 480, (d) primer 1508 and (e) primer 1508 – lower part of the gel.

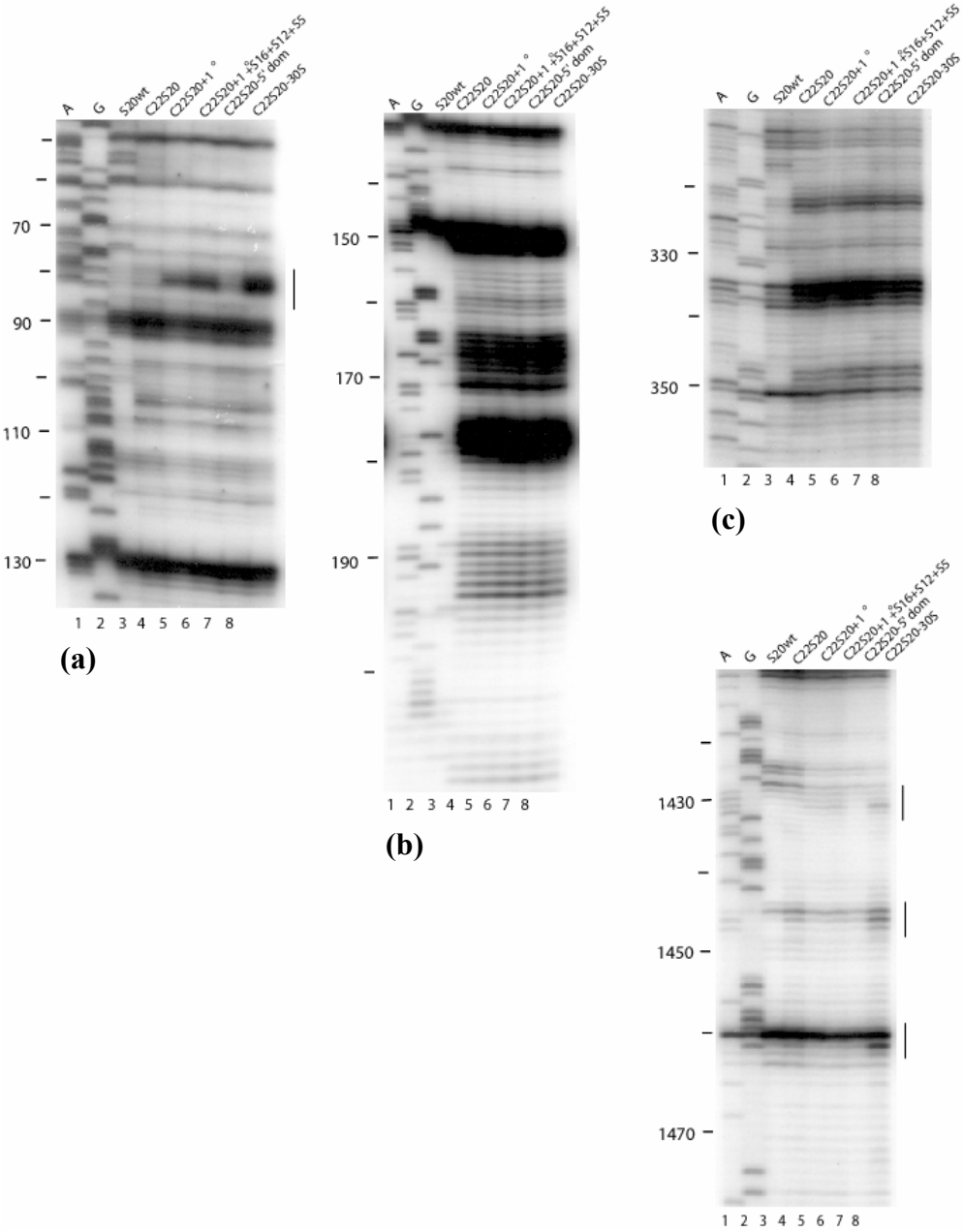
www.manaraa.com



www.manaraa.com

Figure 8. Directed hydroxyl probing of RNPs of different complexities containing Fe(II)-C22S20 analyzed by primer extension. The sites of cleavage were differences between different RNPs are observed are indicated by lines on the side of the gels. The lanes correspond to: sequencing lanes (lanes 1 and 2), and the complexes Fe(II)-C22S20/16S rRNA (lane 3), Fe(II)-C22S20/1^{o}/16S rRNA (1^{o} mixture contains S4, S7, S8, S15 and S17) (lane 4), Fe(II)-C22S20/1^{o}/S12+S16+S5/16S rRNA (lane 5), Fe(II)-C22S20/5'/16S rRNA (5' contains S4, S8, S17, S16, S12, S5) (lane 6) and Fe(II)-C22S20-30S (lane 7). (a) primer 232 – upper part of the gel, (b) primer 232 – lower part of the gel, (c) primer 480, (d) primer 1508.

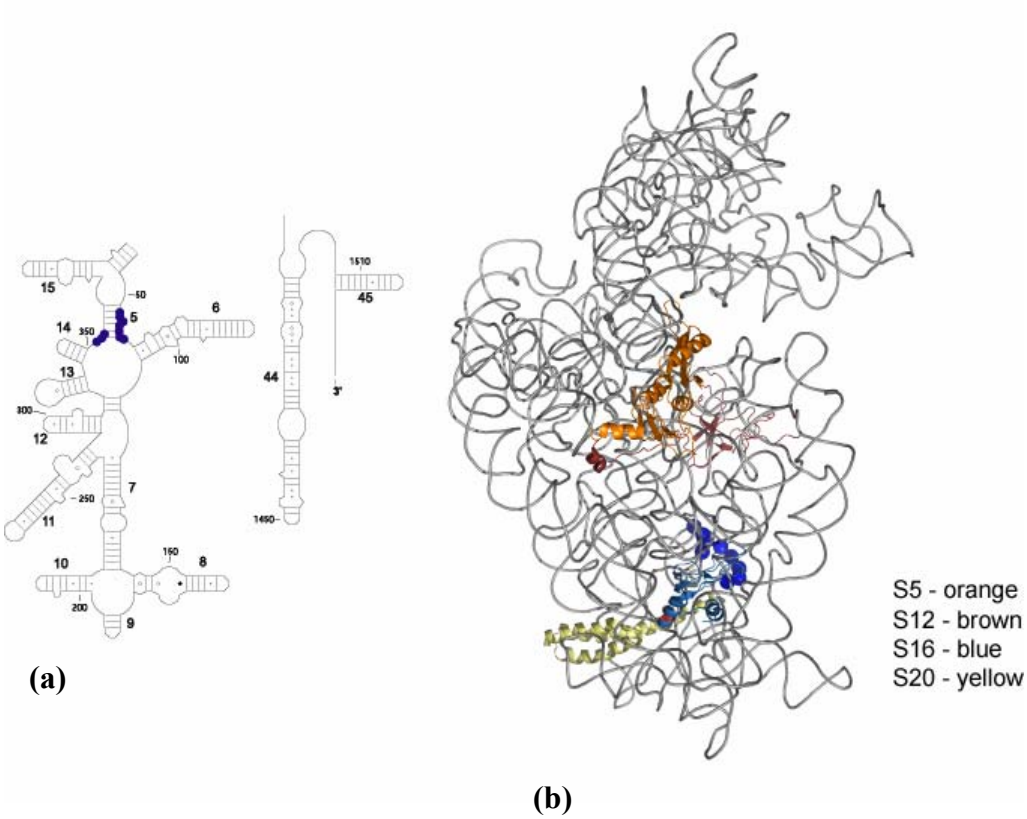
www.manaraa.com



www.manaraa.com

Figure 9. Changes in the intensity of cleavage from C13S20 protein attributed to S16. (a) The blue filled circles indicate nucleotides that become protected from cleavage at assembly of S1, on the secondary structure of 16S rRNA. (b) Three dimensional structure of 16S rRNA shown in gray, from 30S subunit, with S5 orange, S12 brown, S16 blue and S20 yellow. The blue spheres represent the nucleotides indicated on the secondary structure (a).

www.manaraa.com



(a)

(b)

www.manaraa.com

Chapter VIII. RNA-protein interactions in the 30S ribosomal subunit. General summary and future directions

General summary

The complex composition of the 30S ribosomal subunit, 21 ribosomal proteins (r-proteins) and 16S ribosomal RNA (rRNA), makes the study of its assembly a non-trivial problem. Conformational changes in the structure of 16S rRNA are a very important part of the assembly process of the 30S ribosomal subunit.[1-4] In the previous two chapters results obtained by applying two of the methods routinely used in our laboratory to study conformational rearrangements of 16S rRNA during assembly are presented. Both of these methods are using chemical reagents, in particular base-specific probes which modify selectively, like dimethyl sulfate and ketoxal, and hydroxyl radicals which cleave the RNA backbone nonspecifically. Base-specific footprinting of minimal complexes at different temperatures made it possible to distinguish between conformational changes in 16S rRNA that require only the r-protein and the ones that require both the r-protein and heating. Directed hydroxyl radical probing from r-protein S20 of ribonucleoprotein particles (RNPs) of different complexities gave insight into reorientation of the rRNA elements as a result of protein binding.

Base-specific chemical footprinting was used to dissect the temperature dependence of 16S rRNA architecture in individual complexes with the primary r-proteins (S7, S8, S15, S17 and S20). The results published earlier for the sixth primary binding r-protein S4[5] are reanalyzed in the light of the crystal structure of the 30S subunit[6] which is now available, and integrated with our data for the other primary

www.manaraa.com

binding protein/16S rRNA RNPs. As expected, all r-proteins can bind 16S rRNA at low temperature. However not all r-proteins/16S rRNA complexes behave in the same way. As discussed earlier, some RNPs acquire the same conformation regardless of temperature (RNPs containing S17 and S20), others show minor adjustments in 16S rRNA conformation upon heating (RNPs containing S8 and S15), and finally in others 16S rRNA undergoes significant temperature-dependent conformational changes (RNPs containing S4 and S7). Our results correlate very well with the structural and biochemical data available on the 30S subunit. Studies of the *in vitro* assembly of 30S subunit showed that the rate of assembly is strongly temperature dependent,[7] and three distinct stages of assembly have been observed, by using the appropriate temperature regime. The rate determining step is a temperature dependent structural rearrangement of the first reconstitution intermediate, which contains all the primary binding r-proteins. Intermediates similar to those observed *in vitro* have been detected *in vivo*[8] and *s*trains with ribosome biogenesis defects are often cold sensitive.[9] Thus it appears that the temperature-dependent conformational rearrangements of the minimal RNPs that we studied may reflect inherent properties of the 30S subunit assembly *in vitro* and of ribosome biogenesis.

The study of r-protein/16S rRNA binary complexes often gives more detailed information on the RNP, than assembly studies using a full complement of r-proteins, as it was observed in the case of S7, S8 and S15.[10] In the ensemble study at low temperature, the footprints specific for certain r-proteins cannot be observed and the conclusions are not as accurate as when minimal complexes are studied.[10] Also, our approach made it possible to distinguish temperature-dependent stages in the interaction

www.manaraa.com

of 16S rRNA with the r-proteins S4 and S7. Thus our approach makes possible a better understanding of the interaction between 16S rRNA and r-proteins in particular and RNA-protein interactions in general.

Especially intriguing are the conformational rearrangements in 16S rRNA that take place at some distance from the interaction site, which require the presence of both the r-protein and temperature. These long distance effects in 16S rRNA can organize the binding site of other r-proteins that assemble in a sequential manner, modulate interdomain interactions or bring the 16S rRNA into a correct functional conformation. The differential interaction of 16S rRNA with r-proteins illustrates a means for controlling the sequential assembly pathway for complex RNPs and gives insights into some aspects of RNP assembly.

Another approach that proved very useful in the study of conformational rearrangements of 16S rRNA during assembly is directed hydroxyl radical probing.[11] Hydroxyl radicals generated around the probe, Fe(II)-derivatized S20 protein, cleave 16S rRNA and reveal its architecture. An analysis of the cleavage patterns in the minimal complexes and the fully assembled 30S subunit shows intriguing similarities and differences. The study of RNPs of different complexities separately can give information on the assembly, but it is much easier and informative to probe and analyze RNPs of different compositions in the same time. The two probes (Fe(II)-C13S20 and Fe(II)-C22S20) that were used to explore RNPs of different compositions gave useful information and show the versatility of our approach. Interestingly, in some cases the presence of r-proteins that bind at some distance from S20, to other domains of 16S rRNA than the 5' domain influence the cleavage pattern. Roles for different r-proteins

www.manaraa.com

emerge from the preliminary results of this study and future experiments will give a better understanding of the assembly process and of the roles of r-proteins.

Both base-specific footprinting and directed hydroxyl probing illustrate the complexity of the 30S subunit assembly process and the importance of r-proteins in articulating this process. Assembly of r-proteins influences not only the immediate surroundings, but also long distance rearrangements of 16S rRNA that sometime involve simultaneous fulfillment of multiple requirements. Formation of functional small ribosomal subunits involves a series of consecutive events that are controlled by different means, as the presence of a certain protein or the right temperature.

Future directions

Temperature-dependent rearrangements in RNPs containing 16S rRNA. Only binary RNPs containing each of the six primary r-proteins were investigated at different temperatures. There are many possible RNPs containing 16S rRNA that can be analyzed by base-specific chemical footprinting to obtain a more detailed picture of the changes that take place in the architecture of 16S rRNA. For example RNPs containing S16, a secondary binding r-protein are likely candidates for this type of studies. As mentioned in chapter VI, it was shown that S16 plays a very important role in the heat activated transition of the first reconstitution intermediate (RI) to the second reconstitution intermediate RI*.[12] Assembly of S16 requires prior binding of primary r-proteins S4 and S20[13], and S16 has footprints[14] and direct contacts[6] with 16S rRNA both in the 5' and central domain. The binary RNPs containing S4 or S20 along with 16S rRNA behave very differently. In the S20/16S rRNA complex 16S rRNA attains the same conformation at low or high temperature, while for S4, the two conformations are very

www.manaraa.com

different. How will RNPs containing different combinations of the three r-proteins behave? Does S16 bind sequentially to the 5' and central domain? Is binding of S16 to the two different domains controlled by temperature, by the presence of one or both of the primary r-protein or it requires a combination of factors?

Exploring the assembly of 30S subunit with directed hydroxyl radical probing. A complete picture of the 16S rRNA rearrangements during assembly around S20, in the 5' domain, will be obtained by performing probing of RNPs of different complexities with Fe(II)C47S20 and Fe(II)C55S20. But, there are other r-proteins that can be used as probes and may reveal changes that take place in other domains of 16S rRNA. S15 was used to probe the assembly of the central domain,[3,15] and the other primary r-proteins can be used to examine the assembly of the other domains. R-proteins S4 and S7 are especially important because they are considered assembly initiators,[16] and S7 is the only primary r-protein that binds to the 3' major domain. S8 may play an important role in the alignment of the platform (central domain) and the body (5' domain). Probing from secondary binding r-proteins like S5, S16 and S12 that bind multiple 16S rRNA domains[6] can reveal more details on the importance of interdomain alignment during the assembly process.

References

1. Stern, S.; Powers, T.; Changchien, L. M.; Noller, H. F., *Science* **1989,** 244, (4906), 783-90.
2. Williamson, J. R., *Nat. Struct. Biol.* **2000,** 7, (10), 834-7.
3. Jagannathan, I.; Culver, G. M., *J. Mol. Biol.* **2003,** 330, (2), 373-383.
4. Holmes, K. L.; Culver, G. M., *J. Mol. Biol.* **2005,** 354, (2), 340-357.

www.manaraa.com

5. Powers, T.; Noller, H. F., *J. Biol. Chem.* **1995,** 270, (3), 1238-42.

6. Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H., *Science* **2005,** 310, (5749), 827-34.

7. Traub, P.; Nomura, M., *Cold Spring Harb. Symp. Quant. Biol.* **1969,** 34, 63-7.

8. Nashimoto, H.; Held, W.; Kaltschmidt, E.; Nomura, M., *J. Mol. Biol.* **1971,** 62, (1), 121-38.

9. Guthrie, C.; Nashimoto, H.; Nomura, M., *Proc. Natl. Acad. Sci. U.S.A.* **1969,** 63, (2), 384-91.

10. Powers, T.; Daubresse, G.; Noller, H. F., *J. Mol. Biol.* **1993,** 232, (2), 362-74.

11. Culver, G. M.; Noller, H. F., *Methods Enzymol.* **2000,** 318, 461-75.

12. Holmes, K. L.; Culver, G. M., *Nat. Struct. Mol. Biol.* **2004,** 11, (2), 179-186.

13. Mizushima, S.; Nomura, M., *Nature* **1970,** 226, (5252), 1214-18.

14. Stern, S.; Changchien, L. M.; Craven, G. R.; Noller, H. F., *J. Mol. Biol.* **1988,** 200, (2), 291-9.

15. Jagannathan, I.; Culver, G. M., *J. Mol. Biol.* **2004,** 335, (5), 1173-1185.

16. Nowotny, V.; Nierhaus, K. H., *Biochemistry* **1988,** 27, (18), 7051-5.

www.manaraa.com

Acknowledgments

I would like to thank Dr. Gloria Culver for her guidance and support during the time that I have spent in her group. The members of my committee were also very helpful and I thank them for all their contributions. Many other people contributed in different ways to my development during my graduate studies, and I will not be able to mention all of them. Special thanks to Tijana Grove, Nebojsa Milović and Kirthi Narayanaswamy, from whom I have learned many things.

My family and friends were always very supportive and without them this accomplishment would not have been possible. Thank you to all of those that made my life good during these last years.

www.manaraa.com