UNIVERSIDADE DA BEIRA INTERIOR
Ciências

Biorecognition by amino acid-based affinity chromatography for RNA purification

Ana Rita Nunes Martins

Tese para obtenção do Grau de Doutor em

Bioquímica

(3º ciclo de estudos)

Orientadora: Prof. Doutora Fani Pereira de Sousa
Co-orientador: Prof. Doutor João António de Sampaio Rodrigues Queiroz

Covilhã, Outubro de 2013

“Nothing is impossible; the word itself says
‘I'm possible’!”

Audrey Hepburn

Dedication

Esta Tese de Doutoramento é dedicada de forma muito sentida ao meu Pai e à minha Mãe. Ao meu Pai pelo exemplo que sempre foi para mim, pelo seu dinamismo, força de vontade e boa disposição que sempre o acompanharam ao longo da vida, e por me ter ensinado que as palavras “Não” e “Impossível” não existem. À minha Mãe, por me ter incutido a sua generosidade, a sua paciência e a sua forma subtil e desembaraçada de enfrentar a vida.

Todos esses ensinamentos e atitudes definiram a minha personalidade e fizeram-me encarar a vida com uma motivação adicional e com uma perspectiva lutadora, o que me ajudou a alcançar sempre todos os objectivos a que me propus, como é o exemplo desta Tese de Doutoramento.

This work was financed by the Portuguese Foundation for Science and Technology (SFRH/BD/64100/2009) under the program QREN - POPH - Type 4.1 – Advanced Training, co-founded by the European Social Fund and by national funds from the MCTES and also under a research and development project (PTDC/EBB-BIO/114320/2009) and strategic programs (PEst-C/SAU/UI0709/2011) co-founded by the operational program factors of competitiveness – COMPETE and (EXPL/BBB-BIO/1056/2012) co-founded by FEDER funds (FCOMP-01-0124-FEDER) of COMPETE.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Acknowledgments

I wish to thank all those who directly or indirectly contributed for the completion of this doctoral thesis, including those who I will not specifically acknowledge here, but that were also very important in realizing this goal.

First of all, I would like to express my gratitude to my supervisors. To my supervisor Prof. Fani Sousa for her professionalism, dedication and all knowledge provided, for believing in me, for all encouragement, support, and especially for her friendship, my very thank you! To my co-supervisor Prof. João Queiroz, my sincerely appreciation for his geniality and inspiration, for his friendship and availability, as well as the valuable scientific knowledge that he conveyed me, all of which had a remarkable influence in the course of my academic career and in the achievement of this thesis.

I would also like to thanks to Professor Cláudio Maia for his valuable contribution to this work through his expertise in genetics and molecular biology. I am really thankful for the availability and interest he demonstrated for my work through the great scientific discussions and the knowledge transmitted.

To the Portuguese Foundation for Science and Technology (FCT) I want to acknowledge the financial support through my doctoral fellowship (SFRH/BD/64100/2009) that allowed the execution of all my research in Health Science Research Centre at University of Beira Interior (CICS-UBI).

In addition, thanks to CICS-UBI for the host and to the technicians, Ana Sofia Duarte e Margarida Carrilho, for their precious contribution in the organization of my laboratory tasks and for the willingness that they responded to all my requests.

An appreciative note to Professor Luís Passarinha, for his friendship and support, and for allowing me to extend my scientific knowledge and experience by participating in his scientific projects and contributing with my slight experience for the formation of other colleagues.

Moreover, an enormous thanks to my laboratory colleagues for the good working environment, helpfulness and all the excellent moments of scientific discussions but also of fun. I wish them very success in their research projects and their future lives. It was a pleasure working with all of them. However, I could not let to name those who, in addition to co-workers, have become great friends. Indistinctly, my deeply thankful to Ana Martinho, Ângela Sousa, Catherine Nunes, Diogo Maio, Eduarda Coutinho, Fátima Santos, Filomena Silva, Marta Silva and Patrícia Pereira.

To you, Miguel, I am deeply grateful for all the affection, caring, unconditional support and patience throughout these months. You entered my life in the most difficult, demanding and delicate phase of my existence, by coincidence or fate (I do not know!!), but you really demonstrated to be a true companion for life.

Por último, gostaria de agradecer de uma forma muito especial e sentida à minha família pelo amor, confiança, estímulo e apoio incondicional que sempre me deram. Aos meus Pais fica a sentida dedicação desta Tese de Doutoramento. To my sister, I express my thanks for all her caring and for being present at every stage of my life. The complicity that we still have, even at distance, continues to keep us together and encouraged me to move forward, knowing that she will always be there for me. E para o meu cunhado preferido, deixo o meu Obrigado pelos momentos divertidos que em muito me ajudaram a manter o vigor mental.

Resumo

A sequenciação completa do genoma humano ofereceu novos horizontes relativamente à prevenção, diagnóstico e tratamento de doenças humanas. Para além disso, com o avanço da engenharia genética, têm surgido novas tecnologias terapêuticas, entre as quais se destacam a terapia génica utilizando ácidos nucleicos. Apesar destas estratégias génicas se terem iniciado com o DNA (ácido desoxirribonucleico), estudos recentes têm avaliado o potencial interesse terapêutico do RNA (ácido ribonucleico).

O RNA foi recentemente reconhecido como uma molécula fundamental nos processos celulares, com implicações fundamentais na evolução dos organismos, na hereditariedade e na regulação de vários genes, o que destacou o seu vasto potencial terapêutico e conduziu ao aparecimento de várias terapias baseadas em moléculas de RNA. Os resultados promissores dessas novas abordagens terapêuticas têm vindo a reforçar a investigação relacionada com as moléculas de RNA, a avaliar pelo número, cada vez mais elevado, de estudos estruturais, biofísicos e biomédicos presentes na literatura. Além disso, a indústria biotecnológica e farmacêutica começam a visar as moléculas de RNA como uma nova classe de produtos bioterapêuticos.

Um requisito fundamental em todos esses estudos é a obtenção de grandes quantidades de RNA isolado e puro e de integridade assegurada. Como por exemplo, em biologia molecular a purificação de RNA é o primeiro passo chave para avaliar a expressão de um gene, uma vez que a realização bem como a reprodutibilidade e relevância biológica dessa experiência está dependente da quantidade e qualidade das preparações de RNA. Por outro lado, as terapias promissoras e revolucionárias baseadas em RNA, como a vacinação ou utilização de biofármacos recombinantes envolvem formulações de RNA que devem satisfazer critérios de qualidade rigorosos recomendados por agências reguladoras internacionais. No entanto, o processo de purificação das moléculas de RNA pode ser um passo bastante limitante para o sucesso da sua aplicação terapêutica.

O RNA tem uma série de características químicas únicas que se reflectem na sua conformação estrutural. Apesar do RNA ser uma molécula de cadeia simples na sua base, ele tem uma elevada propensão para formar estruturas secundárias e terciárias bastante complexas. São estas compactações próprias das moléculas de RNA que definem as suas importantes funções biológicas. A elevada reactividade química é outra característica própria do RNA e com grande relevância biológica a nível regulatório, por proporcionar mais instabilidade à molécula aumentando a sua susceptibilidade à degradação. Num contexto laboratorial, estas características moleculares do RNA são enormes desafios para a sua extracção e purificação,

pois a sua actividade biológica e integridade podem ser facilmente comprometidas durante os procedimentos devido à presença ubíqua de enzimas que o degradam.

Várias técnicas têm sido desenvolvidas para superar os desafios inerentes ao isolamento e purificação de moléculas de RNA, tais como a extracção com fenol e clorofórmio ou as extracções em fase sólida empregando colunas ou esferas de sílica (SPE), bem como algumas técnicas de purificação com cromatografia líquida de alta eficiência (HPLC) de fase reversa e troca iónica. No entanto, estas ainda apresentam várias limitações nomeadamente em relação ao elevado tempo despendido e à necessidade do uso de solventes tóxicos e condições desnaturantes durante os procedimentos. Por todos estes motivos, torna-se evidente o crescente interesse na avaliação e melhoria das metodologias actualmente utilizadas para o isolamento e purificação de RNA de modo a satisfazer os requisitos necessários à sua aplicação.

A cromatografia é um dos métodos mais diversos e potentes em biotecnologia, tanto a nível analítico como preparativo devido à sua simplicidade, robustez, versatilidade e alta reprodutibilidade. Por sua vez, a cromatografia de afinidade é reconhecida como uma técnica poderosa apresentando grande aplicabilidade na purificação de muitas biomoléculas, incluindo o DNA plasmídico (pDNA) e proteínas, porque explora o reconhecimento biomolecular, ou seja, a capacidade de uma macromolécula biologicamente activa formar complexos específicos e reversíveis com ligandos de afinidade.

Posto isto, o trabalho desenvolvido no âmbito desta tese incide na crescente necessidade de desenvolver novas estratégias de isolamento e purificação para moléculas de RNA de modo a ultrapassar as limitações ainda existentes nas metodologias actuais, contribuindo para a evolução e sucesso da investigação e aplicações terapêuticas do RNA. Para isso, a potencialidade da cromatografia de afinidade foi considerada nesta temática e foi explorada a aplicação de aminoácidos como ligandos de afinidade. Este trabalho foi baseado em vários estudos de reconhecimento molecular e atómico que descrevem a existência de diferentes interacções entre proteínas e ácidos nucleicos nos sistemas biológicos, especialmente com aminoácidos básicos como a histidina e a arginina, e também na hipótese de existirem interacções preferenciais entre os aminoácidos e as bases nucleotídicas. Além disso, estudos recentes de processos de purificação mostraram grande aplicabilidade destes aminoácidos em isolar e purificar moléculas de pDNA biologicamente activas para aplicação em terapia génica e vacinas.

Assim sendo, novas metodologias preparativas e analíticas foram alcançadas ao longo deste trabalho, ou seja, foi possível obter preparações de RNA a partir de diferentes fontes biológicas e de reacções sintéticas com elevados rendimentos e grau de pureza e de integridade preservada, bem como também foi possível o desenvolvimento de um método analítico para a sua quantificação e monitorização.

A potencial aplicabilidade da cromatografia de afinidade, usando histidina como ligando de afinidade, na purificação de moléculas de RNA foi demonstrada pela primeira vez com a purificação do RNA 6S, um RNA regulatório não codificante presente nos procariotas, nomeadamente em *Escherichia coli* (*E. coli*), que tem uma função reguladora importante no processo de transcrição deste organismo. Nas estratégias de purificação com histidina foram usados gradientes de sulfato de amónio devido à presença do anel de imidazol na cadeia lateral aromática do aminoácido, o que permitiria explorar interacções maioritariamente hidrofóbicas entre o RNA e a matriz, quer por interacção com o anel ou por pontes de hidrogénio. Assim, foi utilizado um gradiente decrescente em concentração de sulfato de amónio em três etapas que revelou um reconhecimento bioespecífico com RNA 6S, permitindo a sua purificação de uma mistura complexa de outras moléculas de RNA de baixo peso molecular (sRNA).

Uma segunda nova estratégia desenvolvida com a matriz de histidina permitiu obter simultaneamente o isolamento e purificação das classes de sRNA e RNA ribossómico (rRNA) a partir de lisados celulares de *E. coli*. Neste estudo, ambas as classes foram separadas das impurezas do hospedeiro (DNA genómico e proteínas) com elevado rendimento e grau de pureza quando comparadas com amostras preparadas com métodos de isolamento convencionais baseados na extracção com fenol e clorofórmio. No entanto, a metodologia baseada na cromatografia com histidina demonstrou a vantagem de evitar o uso de produtos químicos tóxicos durante o processo.

A versatilidade da matriz de histidina na purificação de RNA, tanto de moléculas específicas (RNA 6S) como das classes (sRNA e rRNA) sugeriu que o mecanismo de interacção envolveria não só interacções hidrofóbicas, mas também um bioreconhecimento das bases do RNA pela histidina. Apesar das preparações de RNA obtidas com estes métodos necessitarem de caracterização funcional adicional para provar a sua aplicabilidade, o uso da cromatografia de afinidade com histidina representa um grande avanço no isolamento de moléculas de RNA, uma vez que as técnicas tradicionais não têm a capacidade de fraccionar RNA ao nível de um tipo de molécula ou de isolar sRNA e rRNA num só processo. Todavia, a necessidade de aplicar elevadas concentrações de sal nestas metodologias pode ser visto como uma desvantagem, principalmente no que diz respeito à aplicação biotecnológica, pois acarreta maiores custos de processo e tem maior impacto ambiental.

O uso de arginina como aminoácido imobilizado na cromatografia de afinidade foi utilizado na perspectiva de melhorar as técnicas anteriores uma vez que este aminoácido se apresenta sobretudo carregado positivamente, pelo que poderiam ser exploradas interacções electrostáticas para a purificação de RNA utilizando condições de eluição moderadas. Além disso, as interacções de arginina têm sido reconhecidas como as mais prevalentes nos complexos RNA-proteína, aumentando a potencialidade de purificação por um maior bioreconhecimento entre as moléculas de RNA e a matriz de arginina. Deste modo, um

gradiente com o aumento gradual da concentração de cloreto de sódio permitiu o isolamento e purificação do RNA total de extractos celulares eucarióticas. O suporte de arginina demonstrou uma aplicabilidade excepcional para interagir com todas as classes funcionais do RNA apesar da sua diversidade estrutural e as suas diferentes conformações em condições nativas. Essas interacções mais fortes e selectivas parecem advir da cadeia lateral do aminoácido de arginina que apresenta uma multiplicidade para interacções, podendo promover um contacto múltiplo com o RNA, através da sua estrutura açúcar-fosfato ou com as suas bases, atendendo ao seu estado conformacional. Embora as interacções electrostáticas entre os grupos fosfato do RNA e os ligandos de arginina possam desempenhar uma função importante na retenção do RNA na coluna, as interacções com as bases também estão envolvidas e modulam de alguma forma a interacção favorecida e a especificidade encontradas na cromatografia com arginina. Assim, neste processo de purificação verificou-se um elevado rendimento de recuperação do RNA total e pelas análises de controlo de qualidade efectuadas mostrou-se que este apresentava uma elevada integridade bem como uma boa pureza, demonstrada pela difícil detecção de proteínas nas amostras purificadas e pela redução de DNA genómico para concentrações residuais. A eficiência desta técnica de purificação e a aplicabilidade do RNA por ela obtido foi demonstrada num procedimento habitualmente usado em biologia molecular para a análise de expressão génica. As amostras de RNA total foram usadas com sucesso como moldes na reacção em cadeia da polimerase em tempo real (qPCR) para a avaliação da expressão de dois genes controlo geralmente empregues nestes procedimentos.

Reunindo os resultados anteriormente descritos, o isolamento e purificação de RNA de amostras biológicas complexas usando as técnicas de cromatografia de afinidade com aminoácidos imobilizados demostrou vários benefícios em relação a métodos de isolamento actualmente empregues, como a extracção de fenol e clorofórmio ou de SPE, uma vez que simplificam a integração do processo e minimizam o manuseamento das amostras, tornando a cromatografia baseada em aminoácidos útil no desenvolvimento de metodologias em condições não desnaturantes, livres de RNases e solventes orgânicos, particularmente importantes em vários estudos estruturais e funcionais bem como de aplicabilidade terapêutica.

Com os resultados positivos destas metodologias de afinidade na purificação de RNA foi também desenvolvido e validado (de acordo com a legislação internacional e europeia para métodos bioanalíticos) um método analítico para a quantificação e monotorização de RNA. Com a crescente importância, a nível terapêutico, do desenvolvimento de novas ferramentas analíticas para avaliar o RNA, uma vez que ainda existem várias lacunas nas técnicas actuais de quantificação de RNA, tais como a falta de selectividade, tornou-se imperativo avaliar a potencialidade dos suportes de afinidade neste campo. A versatilidade da metodologia foi demonstrada pela sua aplicabilidade na quantificação de RNA de diferentes fontes

eucarióticas e também em amostras complexas de RNA quimicamente sintetizado, o que demonstrou a usa utilidade em múltiplas áreas de investigação.

Desenvolveu-se ainda mais um estudo no isolamento e purificação de RNA com base nas matrizes de aminoácidos, em particular com matriz de arginina, de modo a aproximar a aplicabilidade desta técnica à prática terapêutica. Neste caso, o novo objectivo foi explorar o seu potencial na purificação de moléculas de RNA mensageiro (mRNA) não a partir de células, mas a partir de reacções sintéticas de transcrição *in vitro*, com o interesse de aplicar as moléculas em terapias de vacinação com mRNA no cancro do colo do útero. As moléculas de mRNA que codificavam para as proteínas E6 e E7 do vírus do papiloma humano (VPH) 16 foram purificadas com sucesso de uma série de impurezas próprias das reacções de transcrição *in vitro,* com são o pDNA molde, as enzimas, nucleótidos, sais e tampões, e a sua caracterização em termos de rendimento, pureza e integridade foi avaliada. Neste trabalho, a cromatografia de arginina voltou a demonstrar a sua capacidade singular em melhorar os processos de purificação, pelas vantagens de eliminar passos adicionais e melhorar a economia global do processo de produção.

Em suma, esta tese permitiu o desenvolvimento de novas metodologias para a purificação e quantificação de RNA revelando várias características interessantes das moléculas, incluindo o seu comportamento cromatográfico e interacções naturais que podem ocorrer com os suportes de aminoácidos. Por conseguinte, estes métodos mostraram uma potencial aplicabilidade polivalente, contribuindo para o progresso das investigações fundamentais e terapêuticas baseadas no RNA, suportando a utilização da cromatografia de afinidade baseada em aminoácidos no desenvolvimento futuro de novos processos de preparação de RNA.

Palavras-chave

Arginina, bioreconhecimento, cromatografia de afinidade, histidina, purificação, RNA.

Abstract

Following the decoding of the human genome, a new era was opened for developing new gene therapy strategies employing nucleic acids. Recently, RNA was renowned a central molecule in cellular processes with implications in many diseases as well as in understanding of evolution, becoming one of the most exciting research areas of molecular biology. From basic to applied research, many procedures employ pure and intact RNA molecules. On one hand, RNA purification is a first critical step of a number of molecular biology procedures and its quality is crucial to ensure reproducibility and biological relevance of an experiment. On the other hand, the promising and revolutionary RNA-based therapies of RNA vaccination, gene therapy or recombinant biopharmaceuticals involves RNA formulations which should fulfill rigorous quality criteria recommended by international regulatory agencies. However, the isolation and purification of RNA are critical steps because of the easy degradability of RNA, which can impair chemical stability and biological functionality essential for analysis. Many techniques have been development to overcome the challenges of purifying RNA molecules; nonetheless they still have several limitations in regard to time demanding and the requirement of toxic solvents and denaturing conditions. Therefore, there is a growing demand for the evaluation and improvement of the methodologies currently used for RNA isolation and purification.

Chromatography is undoubtedly one of the most diverse and potent methods in biotechnology, both at analytical, preparative and industrial level due to its simplicity, robustness, versatility and high reproducibility. Affinity chromatography is recognized as a powerful technique with great applicability in the purification of many biomolecules, including plasmid DNA and proteins because it exploits the principle of biomolecular recognition. The work that we have been developing considers new chromatographic strategies for RNA purification, using amino acids as affinity ligands. These studies are based on the fact that many different interactions exist between proteins and nucleic acids in biological systems, involving in particular basic amino acids such as histidine or arginine. New methodologies were accomplished that allowed obtaining RNA preparations from different sources with high recovery yields, purity and integrity. A new analytical method for RNA quantification was also developed in this work.

The applicability of histidine-based affinity chromatography in the purification of RNA molecules was first demonstrated in the separation of 6S RNA, a regulatory non-coding RNA of the prokaryotic *Escherichia coli* (*E.coli*). A specific recognition between the histidine support and 6S RNA allowed its selective purification from a complex mixture of other small RNAs (sRNA). In another strategy, the simultaneous isolation of sRNA and ribosomal RNA from *E.coli* cell lysates, eliminating host DNA and proteins, was also attained by a histidine

chromatographic-based method. Furthermore, arginine matrix was employed in RNA purification from eukaryotic cells demonstrating an exceptional ability to interact with all functional classes of RNA, despite their structural diversity and different folding states, enabling their isolation from impurities of eukaryotic crude cell extracts. Moreover, an analytical technique based on arginine affinity support for quantification and quality assessment of total RNA from different eukaryotic cells and synthetic RNA samples was also developed and validated, according to international and European legislation for bioanalytical methods.

More efforts into RNA purification were developed with amino acid-based matrices, in particular with arginine-agarose matrix, in order to approach this technique to therapeutic application of RNA. The new goal was to exploit its applicability in purifying messenger RNA (mRNA) molecules not from cells, but from synthetic crudes of *in vitro* transcription reactions, pursuing mRNA vaccination for cervical cancer. In this work, arginine-based chromatography also showed its singular capability to improve purification processes, showing the advantages of eliminating additional steps and improving global economics of the production process.

The development of these new methodologies revealed several interesting characteristics of RNA molecules, including their chromatographic behavior and natural interactions that can occur between amino acids-based supports and RNA molecules. Accordingly, these methods demonstrated a potential multipurpose applicability by aiding in molecular biology RNA-based analysis and RNA therapeutics, which support the interest in applying amino acid-based affinity chromatography for the future development of new RNA isolation, purification and quantification processes.

Keywords

Affinity chromatography, arginine, biorecognition, histidine, purification, RNA.

Thesis Overview

This thesis is structured in four main chapters.

The **first chapter** includes an introduction that is divided in two sections. One section explains the interest on RNA molecules to be potentially applied in novel therapies such as gene therapy and RNA vaccination and the second section presents a brief explanation about the relevance of purifying RNA molecules, as well as the main challenges and concerns regarding RNA purification. In addition, a brief review of the literature related to the techniques used in RNA purification regarding their advantages and disadvantages is discussed together with the improvements that have been done lately. This second section is presented as a publisher paper review form (Paper I).

The **second chapter** presents the main purpose and the specific goals that were established for the development of this research work.

In the **third chapter**, the results obtained during this work are presented and discussed in the form of original research papers organized as follows:

Paper II - A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography

Paper III - Histidine affinity chromatography-based methodology for the simultaneous isolation of *Escherichia coli* small and ribosomal RNA

Paper IV - A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography

Paper V - New approach in RNA quantification using arginine-affinity chromatography: potential application in eukaryotic and chemically synthesized RNA

Paper VI - Arginine-affinity chromatography for mRNA vaccines purification

Finally, the **fourth chapter** outlines the concluding remarks about this work, regarding the initial hypothesis of using affinity chromatography with immobilized amino acids in the purification of RNA molecules. Furthermore, some future work are suggested to complement the important findings achieve in this study.

Index

List of Figures	xxvii
List of Tables	xxix
List of Scientific publications	xxxi
List of Scientific communications	xxxv
Chapter 1	1
1. Gene therapy and vaccination with RNA	3
1.1. Introduction	3
1.2. Overview of gene therapy	4
1.3. RNA-based therapies	11
1.3.1. Catalytic RNAs: Ribozymes	11
1.3.2. Antisense oligoribonucleotides	13
1.3.3. Agents of RNA interference	13
1.3.4. Aptamers	16
1.3.5. mRNA vaccination	17
References	19
2. Current issues in RNA preparation: approaching affinity chromatography into RNA purification challenges (Paper I)	23
Chapter 2	67
Aims of the thesis	69
Chapter 3	73
Paper II. A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography	75
Paper III. Histidine affinity chromatography-based methodology for the simultaneous isolation of *Escherichia coli* small and ribosomal RNA	85
Paper IV. A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography	97
Paper V. New approach in RNA quantification using arginine-affinity chromatography: potential application in eukaryotic and chemically synthesized RNA	111
Paper VI. Arginine-affinity chromatography for mRNA vaccines purification	125
Chapter 4	151
Concluding remarks	153
Future perspectives	157

List of Figures

Figure 1 – Compiled data on gene therapy clinical trials provided by regulatory agencies 5

Figure 2 – The *in vivo* and *ex vivo* paths used in gene therapy 6

Figure 3 – Vector systems commonly used in gene therapy clinical trials 7

Figure 4 – Ribozymes activities 12

Figure 5 – miRNAs, shRNAs and siRNAs pathways for RNAi in mammalian cells 15

Figure 6 - Treatment of cancer patients with tumour RNA-transfected DCs 18

List of Tables

Table 1 - Viral vectors use for gene therapy 7

Table 2 – Non-viral delivery systems used in gene therapy approaches 9

List of Scientific publications

Papers related to this Thesis

- **Current issues in RNA preparation: approaching affinity chromatography into RNA purification challenges**
  R. Martins, J. A. Queiroz, F. Sousa
  Submitted for publication (2013)

- **Arginine-affinity chromatography for mRNA vaccines purification**
  R. Martins, C. J. Maia, J. A. Queiroz, F. Sousa
  To submit (2013)

- **New approach in RNA quantification using arginine-affinity chromatography: potential application in eukaryotic and chemically synthesized RNA**
  R. Martins, J. A. Queiroz, F. Sousa
  Analytical and Bioanalytical Chemistry. 2013. 405(27): 8849-8858

- **A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography**
  R. Martins, C. J. Maia, J. A. Queiroz, F. Sousa
  Journal of Separation Science. 2012. 35(22): 3217-3226

- **Histidine affinity chromatography-based methodology for the simultaneous isolation of *Escherichia coli* small and ribosomal RNA**
  R. Martins, J. A. Queiroz, F. Sousa
  Biomedical chromatography. 2012. 26(7): 781-788

- **A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography**
  R. Martins, J. A. Queiroz, F. Sousa
  Journal of Molecular Recognition. 2010. 23(6): 519-524

Papers not related to this Thesis

- **Performance of hydrophobic interaction ligands for human membrane-bound catechol-O-methyltransferase purification**
  F. M. Santos, A. Q. Pedro, R. F. Soares, R. Martins, M. J. Bonifácio, J. A. Queiroz, L. A. Passarinha
  Journal Separation Sciences. 2013. DOI: 10.1002/jssc.201300010

- **Screening of gellan gum as an ionic and hydrophobic chromatographic matrix for biomolecule purification**
  L. A. Rocha, A. Gonçalves, F. Silva, R. Martins, A. Sousa, L. A. Passarinha
  Submitted for publication (2013)

- **Matriz cromatográfica baseada no polímero polissacárido gelana**
  L. A. Rocha, F. M. Santos, F. Silva, R. Martins, A. Sousa, L. A. Passarinha
  Portuguese Patent 106446. Sep 7, 2012

- **Characterization of polyplexes involving small RNA**
  P. Pereira, A. F. Jorge, R. Martins, A. A. Pais, F. Sousa, A. Figueiras
  Journal Colloid Interface Science. 2012. 387 (1):84-94

List of Scientific communications

Oral communications related to this Thesis

- **Affinity-based method for RNA purification pursuing mRNA vaccination**
  R. Martins, C. J. Maia, J. A. Queiroz, F. Sousa
  19th ISSS - International Symposium on Separation Sciences: New achievements in Chromatography 2013. Poreč, Croatia

- **A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography.**
  R. Martins, C. J. Maia, J. A. Queiroz, F. Sousa
  32th ISPPP - International Symposium on the Separation of Proteins, Peptides and Polynucleotides 2012. Istanbul, Turkey

- **A new effective method for purifying *Escherichia coli* small and ribosomal RNA using histidine affinity chromatography**
  R. Martins, J. A. Queiroz, F. Sousa
  Affinity 2011. Tavira, Portugal

- **RNA purification by histidine affinity chromatography**
  R. Martins, J. A. Queiroz, F. Sousa
  IV National Meeting of Biochemistry Students 2009. Covilhã, Portugal

Oral communications not related to this Thesis

- **Evaluation of human membrane-bound catechol-O-methyltransferase purification by hydrophobic interaction chromatography**
  F. M. Santos, A. Q. Pedro, R. F. Soares, R. Martins, M. J. Bonifacio, J. A. Queiroz, L. A. Passarinha
  32th ISPPP - International Symposium on the Separation of Proteins, Peptides and Polynucleotides 2012. Istanbul, Turkey

Poster communications related to this Thesis

- **Arginine based-chromatography as a new approach for prokaryotic and eukaryotic RNA isolation**
  R. Martins, J. A. Queiroz, F. Sousa
  7º National Meeting of Chromatography 2012. Porto, Portugal

- **Isolation of RNA from cell lysates using histidine affinity chromatography. International**
  R. Martins, J. A. Queiroz, F. Sousa
  30th ISPPP - Symposium on the Separation of Proteins, Peptides and Polynucleotides 2010. Bologna, Italy

- *Ribosomal RNA isolation from cell lysates by histidine affinity chromatography*
  R. Martins, J. A. Queiroz, F. Sousa
  12th SBCN - International meeting and workshop of the Society for Biochromatography and Nanoseparations 2010. Lyon, France

- **A new affinity approach to isolate RNA species with histidine-chromatography**
  R. Martins, J. A. Queiroz, F. Sousa
  Affinity 2009. Reykjavik, Iceland

Poster communications not related to this Thesis

- **Structural and functional characterization of polyplexes for small RNA delivery**
  P. Pereira, A. F. Jorge, R. Martins, A. A. Pais, F. Sousa, A. Figueiras
  8th World Meeting on Pharmaceutics, Biopharmaceutics and Pharmaceutical Technology 2012. Istanbul, Turkey

- **Development of a Gellan Gum stationary phase as a new support for biomolecules purification**
  L. A. Rocha, A. Gonçalves, F. Silva, R. Martins, A. Sousa, L. A. Passarinha
  32th ISPPP - International Symposium on the Separation of Proteins, Peptides and Polynucleotides 2012. Istanbul, Turkey

Chapter 1

1. Gene therapy and vaccination with RNA

1.1. Introduction

In recent decades, the advances in molecular biology combined with the culmination of the decoding of the human genome have provided a genetic understanding of cellular processes and disease pathogenesis. Numerous genes involved in disease have been identified as targets for therapeutic approaches and a new era was opened for developing new gene therapy strategies employing nucleic acids. Although the gene-based therapeutic strategies started to be developed using DNA, a large number of studies are in progress in which the therapeutic potential of RNA is evaluated.

The concept of using RNA molecules as therapeutic agents rose from a variety of newly scientific discoveries that revealed RNA to be a versatile biological macromolecule fundamental in mobilizing and interpreting genetic information and essential in cellular processes of all living systems. The research in this area has been fuelled with the exploitation of the inherent properties of RNAs with the purpose to interfere with or repair dysfunctional nucleic acids or proteins and to stimulate the production of therapeutic gene products in a variety of pathological situations. The simplicity of RNA engineering combined with its versatility in structure and function has highlighted the use of RNA-based strategies for therapy. The first generation of RNA therapeutics is now being evaluated in clinical trials, raising significant interest in this emerging area of medical research.

1.2. Overview of gene therapy

Gene therapy is a highly promising therapeutic method to treat various diseases, including both genetic and acquired disorders. In principle, gene therapy uses genetic information for the treatment or prevention of a disease. It involves the transfer of a therapeutic genetic material into specific cells of an individual in order to repair a defective gene or to introduce a new gene whose function is to cure or to favourably modify the clinical course of a condition (Verma and Weitzman, 2005).

Virtually all cells in the human body contain genes, making them potential targets for gene therapy. Nevertheless, cells can be divided into two major categories: somatic cells (most cells of the body) or cells of the germ-line (eggs or sperm). In theory it is possible to transform either somatic cells or germ cells. However, somatic cells are non-reproductive and therefore somatic cell therapy is viewed as a more conservative, safer approach, because it affects only the targeted cells in the patient and is not passed on to future generations. All gene therapy to date on humans has been directed at somatic cells, whereas germ-line engineering in humans remains controversial and prohibited in for instance the European Union (Wivel, 2002).

Historically, treating diseases by genetic engineering is an original conceptualization of the investigators Avery, MacLeod and McCarthy that pioneered the notion and demonstrated that genes could be transferred within nucleic acids in the early 1940s (Wolff and Lederberg, 1994). Soon after, viruses were envisioned as potential tools for human's benefit, in theoretical studies in somatic-cell genetics or possibly in gene therapy. Viral genomes were then used for the development of the first relatively efficient methods for gene transfer into mammalian cells in culture. In the late 1970s, the discovery of recombinant DNA technology provided the tools to efficiently develop gene therapy. In the decades that followed, tremendous advances in this technology enabled the manipulation of viral genomes, isolation of genes, identification of mutations involved in human diseases, characterization and regulation of gene expression, and engineer various delivery systems. In the early 1990s, the first human gene therapy clinical trial was finally approved for treating a form of immune deficiency called adenosine deaminase deficiency. Within a short time period, gene therapy has moved from the conceptual stage to technology development and laboratory research to clinical translational trials, which is clearly demonstrated by the increased number of gene therapy clinical trials approved since 1989 as well as the widened range of diseases for which gene therapy trials have been approved and the various gene types that have been used over recent years (figure 1).

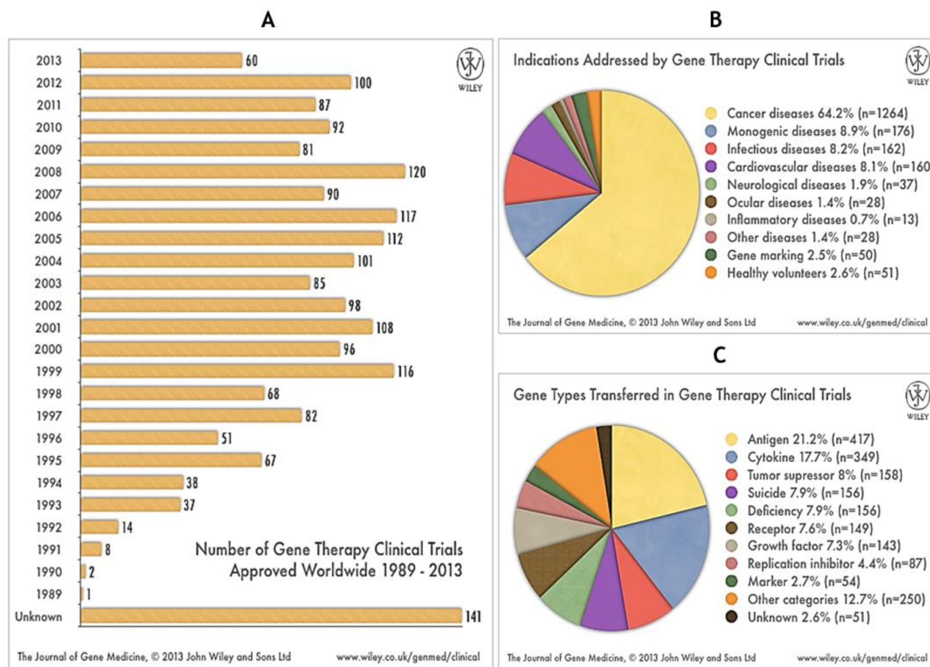

Figure 1 – Compiled data on gene therapy clinical trials provided by regulatory agencies. Gene therapy trials are categorised according to: (A) annual number of trials approved/initiated 1989-2013; (B) diseases addressed and (C) gene types transferred.
(Available at: *www.wiley.com/legacy/wileychi/genmed/clinical/, October 2013*).

Technically, a gene therapy procedure encompasses genetically altering or modifying cells or tissues with a composition of exogenous genetic materials. This composition is an active substance which consists of recombinant nucleic acids used in administration to human beings with a view to regulate, repair, replace, add or delete a genetic sequence, and its therapeutic, prophylactic or diagnostic effect relates directly to the recombinant nucleic acid sequence it contains, or to the product of genetic expression of this sequence (Gascón *et al.*, 2013). This therapeutic product can be introduced and expressed in the cells of a patient by two major general approaches: transfer of genes into patient cells inside the body (*in vivo*), often with the goal of targeting particular tissues (or organs), or outside of the body (*ex vivo*), where the patients' cells are isolated, expanded and modified *ex vivo* before being reintroduced into the same subject (O'Connor and Crystal, 2006). Figure 2 demonstrates the general process of both strategies used in transfer of genes for gene therapy treatments.

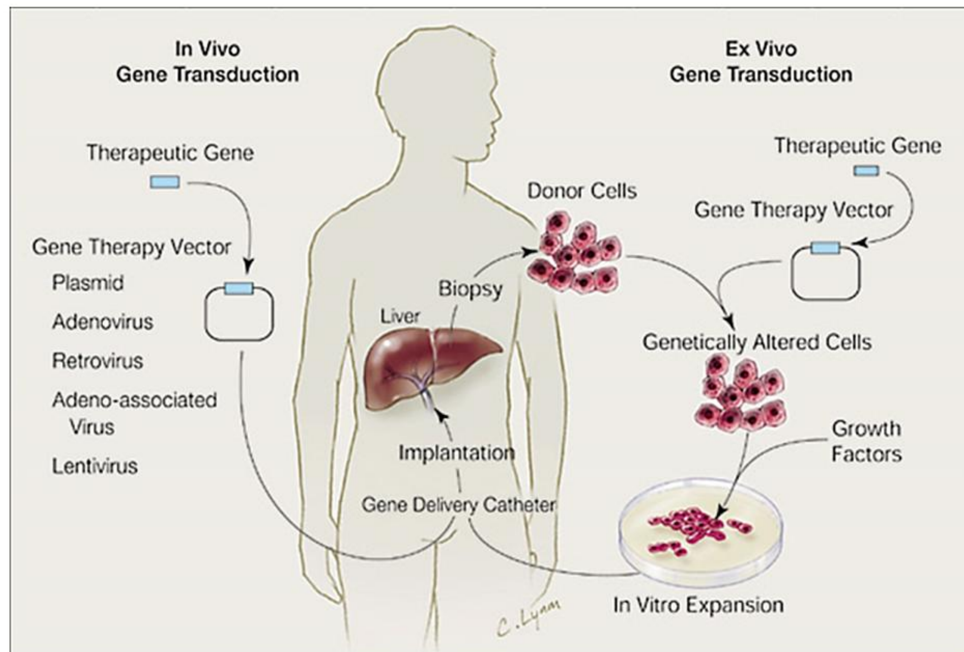

Figure 2 – The *in vivo* and *ex vivo* paths used in gene therapy. The left side of the illustration shows the *in vivo* approach to gene therapy where the therapeutic nucleic acid is directly delivers to the patient. The gene can be delivered into the target cell by several delivery systems, commonly called as gene therapy vectors. On the right side is the representation of the *ex vivo* approach that involves the transfer of genes into cultured cells which were previously isolated from the patient or other donors. These genetically altered cells are proliferated or cultured *in vitro* and subsequently implanted into the patient. Gene transfer *in vitro* can be performed by the same delivery systems as those used in *in vivo*. (Adapted from Kaji and Leiden (2001)).

In general, a therapeutic gene is delivered to the cell using a carrier, or vector, rather than directly inserted into patient's cells, due to the reduced uptake into the cells of naked therapeutic nucleic acids. A key factor in the success of gene therapy is the development of delivery systems that are capable of efficient gene transfer in a variety of tissues, without causing any associated pathogenic effects. A gene transfer system can be considered ideal if the following aspects have been satisfied: specificity and efficiency of gene transfer; magnitude and duration of expression; immunogenicity and manufacturing (Verma and Weitzman, 2005). To make gene transfer more efficient, specific and safe, a variety of different vectors and delivery techniques have been developed and studied, and applied in gene therapy trials (Figure 3). Generally, these methods can be divided into two categories, viral gene delivery and non-viral gene delivery, depending on the vectors involved.

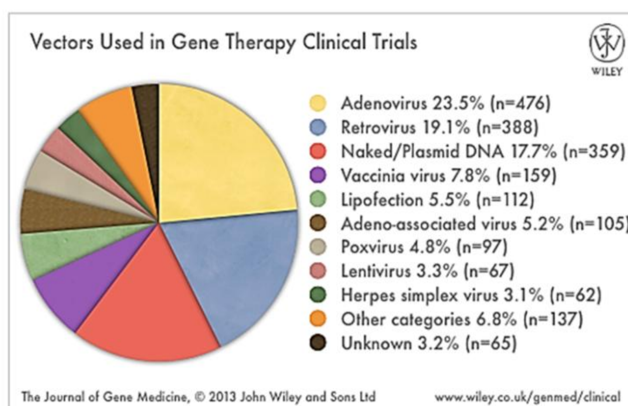

Figure 3 – Vector systems commonly used in gene therapy clinical trials.
(Available at: *www.wiley.com/legacy/wileychi/genmed/clinical/*, *October 2013*).

Currently, the most common type of vectors are viruses that have been genetically altered to carry normal human nucleic acids (Ginn *et al.*, 2013). To date, five main classes of viral vectors have been tested for clinical applications. These include adenoviruses, adeno-associated viruses, retroviruses, lentiviruses and herpes simplex viruses (table 1). Viral vectors are in fact the most effective because they offer higher transduction efficiency and long-term gene expression, but their application can be limited by their immunogenicity, oncogenicity and the small size of the nucleic acids they can transport (Walther and Stein, 2000).

Table 1 - Viral vectors use for gene therapy. (AAV, adeno-associated viruses; dsDNA, double-stranded DNA; ssDNA/RNA, single-stranded DNA/RNA). (Adapted from Sheridan (2011)).

	Adenovirus	AAV	Retrovirus/Lentivirus	Herpesvirus
Family	Adenoviridae	Parvoviridae	Retroviridae	Herpesviridae
Genome	dsDNA	ssDNA	ssRNA+	dsDNA
Infection/tropism	Dividing and non-dividing cells	Dividing and non-dividing cells	Dividing cells	Dividing and non-dividing cells
Host genome interaction	Non-integrating	Non-integrating	Integrating	Non-integrating
Transgene expression	Transient Potential	Long lasting	Long lasting	Potential long lasting
Packaging capacity	7.5 kb	4.5 kb	8 kb	>30 kb

Over the past decade, numerous non-viral methods for gene transfer have been proposed, including physical methods and the employment of chemical vectors (table 2). These non-viral vectors offer several advantages over viral vectors: simplicity of large scale production, low immunogenicity, low toxicity and potential for more tissue specificity. The simplest non-viral gene delivery system uses 'naked' nucleic acids, such as plasmid DNA, which can be delivered into cells or tissues by physical methods such as electroporation, gene gun delivery, sonoporation or hydrodynamic injection. Meanwhile, chemical vectors can be used to encapsulate nucleic acids, exerting a protective effect. Currently in use, include inorganic nanoparticles, as calcium phosphate, and synthetic or natural biodegradable particles such as cationic lipids (forming lipoplexes upon mixing with nucleic acids) or cationic polymers (forming polyplexes upon mixing with nucleic acids) (Gascón *et al.*, 2013). However, despite recent technological advances, the main limitation of non-viral systems is their low transfection efficiency, although it has been improved by different strategies and the efforts are still ongoing (Wang *et al.*, 2013).

Table 2 – Non-viral delivery systems used in gene therapy approaches. (Adapted from Gascón *et al.* (2013)).

Category	System for gene delivery	
Physical methods	Needle injection	
	Balistic DNA injection	
	Electroporation	
	Sonoporation	
	Photoporation	
	Magnetofection	
	Hydroporation	
Inorganic particles	Calcium phosphate	
	Silica	
	Gold	
	Magnetic	
Synthetic or natural biodegradable particles	Polymeric-based vectors	Poly(lactic-co-glycolic acid)
		Poly lactic acid
		Poly(ethylenimine)
		Chitosan
		Dendrimers
		Polymethacrylates
	Cationic lipid-based vectors	Cationic liposomes
		Cationic emulsions
		Solid lipid nanoparticles
		Cationic liposomes
	Peptide-based vectors	Poly-L-lysine
		Other peptides to functionalize other delivery systems: cell-penetrating peptides, protamine.

Gene therapy is a relatively new paradigm in medicine with enormous therapeutic potential. A major motivation for gene therapy has been the need to develop novel treatments for diseases for which there is no effective conventional treatment. As previous stated, the spectrum of gene therapy applications has now broadened considerably to every area of molecular medicine offering new possibilities of mitigating, and even curing, numerous of medical conditions ranging rare inherited monogenic disorders, metabolic diseases, infections and even complex disorders such as cancer (Ginn *et al.*, 2013). The traditional gene therapy was focused on delivery of DNA encoding therapeutic proteins into cells. Depending on purpose and delivery method, successful gene transfer could have several outcomes: to modify defective host genes, to replace deficient host genes, to insert into the host genome or to stay in the nucleus without integration into the host genome. The subsequent transgene expression could restore normal cellular processes or induce new cellular responses. With the research in recent years, current gene therapy is only restricted to deliver DNA. The delivery of any other therapeutic nucleic acid materials, such as RNA who interferes with gene expression by regulating post-transcription or translation, could also be included into the concept of gene therapy (Strachan and Read, 1999 , Wang *et al.*, 2013).

1.3. RNA-based therapies

RNA was once considered to be just an intermediate molecule in taking genetic information from the genome to the ribosome, but that view has been changing rapidly by the recent knowledge coming from basic RNA research. Recently, RNA was renowned as a central molecule in cellular processes and gene regulation. This centrality of RNA reflects its unprecedented biochemical properties. The linear sequence of RNA makes it a simple source of genetic information. The property of RNA to form secondary structure, shielding some sequences while exposing others for recognition, facilitates its interactions with other molecules. In a more complex way, RNA can assume tertiary structures that present surfaces for interactions and contain internal environments that create binding sites for metal ions that can promote catalytic reactions (Sharp, 2009).

It is now clear that RNA is a versatile molecule that play key roles in many important biological processes like splicing, editing, protein export and others, and it can also act catalytically, like enzymes (Soll *et al.*, 2001) which underscore the therapeutic potential of RNA as a new gene therapy tool. RNA-based therapeutics make use of the mechanism of activity of the various RNA molecules, which include catalytically active RNA molecules (ribozymes), inhibitors of mRNA translation (antisense oligoribonucleotides), the agents of RNA interference in gene expression (small RNAs), and RNAs that bind proteins and other molecular ligands (aptamers) (Burnett and Rossi, 2012).

1.3.1. Catalytic RNAs: Ribozymes

The discovery that RNA can act as an enzyme changed the paradigm of the central dogma of molecular biology (Lehman, 2010), and led to the development of a new class of therapeutics based on RNAs enzymes. Ribozymes, or RNA enzymes, are RNA molecules that can mediate their own cleavage or splicing or act as enzymes to promote reactions on substrate RNA molecules (Khan, 2006) (figure 4).

Two types of RNA enzymes - hammerhead and hairpin ribozymes – have been the main focus of efforts towards assessing the potential therapeutic utility of ribozymes. These ribozymes were found to mediate inhibition of gene expression through the binding of messenger RNA (mRNA) by complementarity and inducing its site-specific cleavage (figure 4a) with the particularity that ribozymes recycle themselves ready to repeat this process multiple times. Early preclinical works showed that such RNA enzymes could repeatedly cleave practically any pathogenic transcript, which supported their use as therapeutic tools for manipulation of gene expression (Hauswirth *et al.*, 2001).

Several phase I and II clinical trials have been initiated using *trans*-cleaving ribozymes in a small number of patients with infectious diseases or cancer. In these studies the ribozymes

have been delivered to the patients either by delivery systems or by direct injection of a synthetic ribozyme that contains chemical modifications that greatly increase the ribozyme's stability in biological fluids. However, the gene therapy-based trials have focused upon developing ribozyme-based treatments for individuals infected with the human immunodeficiency virus (HIV) (Burnett and Rossi, 2012).

Recently, the ribozymes with self-splicing ability were exploited to trans-splice RNA targets in order to repair mutant mRNA molecules giving rise to genetic diseases (figure 4b). RNA repair is an alternative way to control gene expression at the mRNA level, repairing mutant mRNAs rather than destroying them. This method uses the splicing and editing processes that create mRNAs from DNA to replace mutant regions in mRNA. The initial studies focused on RNA repair used a trans-splicing version of a group I ribozyme to repair mutant lacZ transcripts in bacteria and mammalian cells. These studies showed that the ribozyme was able to repair the mutant RNA by recognizing the target transcript by base pairing with it, cleaving off mutant sequences and linking a wild-type sequence onto the cleaved product (Phylactou *et al.*, 1998).

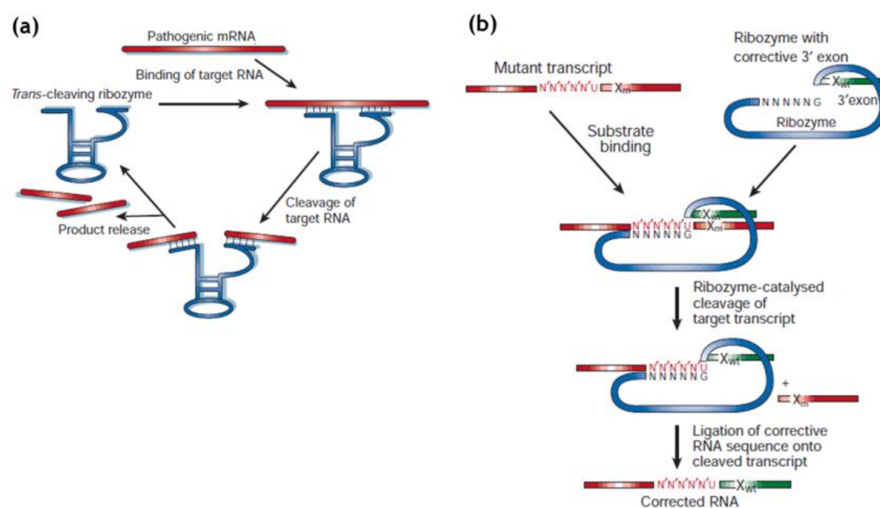

Figure 4 – Ribozymes activities. (a) *Trans*-cleaving ribozymes can bind pathogenic mRNAs through base-pairing interactions and perform sequence-specific cleavage by phosphodiester isomerization, releasing the reaction products. Ribozymes can repeat this process with multiple turnover. (b) *Trans*-splicing ribozymes can repair a mutant RNA by recognizing the target transcript upstream of a mutation site (Xm) by base pairing with it. The mutant sequence is cleaving off and an exon with a wild-type sequence (Xwt) is ligated onto the cleavage product to generate a corrected transcript. (mRNA, messenger RNA). (Adapted from Sullenger and Gilboa (2002)).

1.3.2. Antisense oligoribonucleotides

Antisense RNAs are small, diffusible, highly structured RNAs that act via sequence complementarity on target RNAs called sense RNAs. In eukaryotes, some processes like splicing or editing make use also of complementary small RNAs (Brantl, 2002). Antisense RNA has long been thought of as a promising technique for disease therapy. The concept underlying antisense technology is relatively straightforward: the use of a sequence, complementary by virtue of Watson-Crick base pair hybridization, to a specific mRNA can inhibit its expression and then induce a blockade in the transfer of genetic information from DNA to protein. Conceptual simplicity, the possibility of rational design and relatively inexpensive cost has led to the widespread use of these short fragments of RNA as therapeutic agents (Dias and Stein, 2002). The success on clinical trials of this RNA-based therapy has led to the first antisense RNA approved by the Food and Drug Administration (FDA) for commercialization in 1998. Vitravene (Isis Pharmaceuticals/Novartis) is the brand name for Fomivirsen, an antiviral drug used in the treatment of cytomegalovirus retinitis (CMV) in immune compromised patients, including those with acquired immunodeficiency syndrome. Fomivirsen is a synthetic 21-nucleotide sequence with phosphorothioate linkages (which are resistant to degradation by nucleases), which blocks translation of viral mRNA by binding to the complementary sequence of the mRNA transcribed from the coding segment of a key CMV gene (Grillone and Lanz, 2001). Since then, the significant advances in RNA chemistry led to the creation of second-generation antisense technology that was expected to overcome many of the limitations of the original approaches and expand the use of this technology to other diseases. In fact, this has been successful and in the beginning of this year FDA approved an injectable antisense RNA, called KYNAMRO™ (mipomersen sodium) from Genzyme Company, which is an oligoribonucleotide inhibitor of apolipoprotein B-100 synthesis indicated as an adjunct to lipid-lowering medications and diet to reduce low density lipoprotein-cholesterol, apolipoprotein B, total cholesterol, and non-high density lipoprotein-cholesterol in patients with homozygous familial hypercholesterolemia (Genzyme, 2013).

1.3.3. Agents of RNA interference

The Nobel-prize winning in 1998 for the discovery of the mechanism of activity of a class of small RNA (sRNA) molecules produced by eukaryotes aroused a novel therapeutic approach to treat human diseases, called RNA interference (RNAi) technology. The cellular process of RNAi uses sRNAs to mediate resistance to both endogenous parasitic and exogenous pathogenic nucleic acids, and regulates the expression of protein-coding genes through post-transcriptional gene silencing (PTGS). PTGS is regulated by two distinct mechanisms: sequence-specific cleavage of perfectly complementary mRNAs and translational repression and degradation of mRNAs with imperfect complementarity. Small interfering RNA (siRNA) and microRNA (miRNA) are the biological agents of RNAi, a family of regulatory non-coding

RNAs of 19-28 nucleotides in length (figure 5) (Burnett and Rossi, 2012, Kim, 2005). siRNAs are short double-stranded RNA (dsRNA) with 2 nucleotides overhangs at the 3′- ends. In the cytoplasm, siRNAs are loaded into a protein complex called the RNA-induced silencing complex (RISC). The loaded RISC complex then scans all intracellular mRNA for a target mRNA with a complementary sequence to the loaded siRNA. If a target mRNA is found by the loaded RISC, the target mRNA is cleaved and degraded, successfully inhibiting the translation of the target gene. siRNAs can be generated in several ways. In some cases, long dsRNA is introduced in the cell, either by a virus or endogenous RNA expression (microRNA). The enzyme Dicer cleaves the long duplex RNAs into siRNAs (Guo *et al.*, 2010).

On the other hand, miRNAs are originated from endogenous genome DNA sequence and are first transcribed in the nucleus as parts of long primary miRNA transcripts (referred to pri-miRNA) with 5′- caps and 3′- polyA tails. miRNAs with hairpin structures are then processed into pre-miRNAs by the ribonuclease Drosha. The pre-miRNAs are subsequently transported out of the nucleus to cytoplasm by the dsRNA-binding protein Expotin-5, and processed to mature miRNAs by the endoribonuclease Dicer. Similar to siRNA-mediated silencing pathway, miRNA is then loaded into RISC. However, its mode of action is dependent on the extent of sequence complementarity between the miRNA and the target mRNA. When a miRNA matches the sequence of the mRNA completely, the miRNA/RISC complex mediates the cleavage of the mRNA using the same mechanism as siRNA. For miRNAs that only partially match the mRNA's sequence, the miRNA/RISC complex induces translational inhibition and subsequent mRNA degradation. miRNAs silencing is arguably more complex than siRNA silencing, owing to the fact that miRNAs only require partial sequence complementarity to silence genes (Guo *et al.*, 2010, Lin *et al.*, 2006). The miRNA mechanism is not fully understood and some diseases are suggested to be linked to aberrant miRNA expression and function (Soifer *et al.*, 2007).

For therapeutic purposes, siRNA has been the focused molecule in the RNAi pathway. In fact, siRNAs are even being used to interfere with aberrant miRNAs. PTGS may be efficiently induced by delivering exogenously synthetic siRNA molecules to cells. Chemically synthesized siRNA duplex or siRNA molecules prepared *in vitro* from dsRNAs by incubating with recombinant Dicer protein, are commonly used in research for gene silencing. In this last strategy, Dicer-processed siRNA products simply bypass the Dicer cleavage step. Another way to introduce siRNA into cells is to express short hairpin RNAs (shRNA) genes in plasmid vectors (Guo *et al.*, 2010). shRNA is a sequence of RNA which is created in the cell from a DNA construct encoding a sequence of single stranded RNA and its complement, separated by a stuffer fragment, allowing the RNA molecule to fold back on itself, creating a dsRNA molecule with a hairpin loop. In the cell, it is transcribed under the control of RNA Polymerase-II or Polymerase-III promoters and folds into a structure resembling a siRNA duplex. shRNAs are then processed by Dicer into siRNAs (figure 5) (Guo *et al.*, 2010, Tokatlian and Segura, 2010, Xiang *et al.*, 2006).

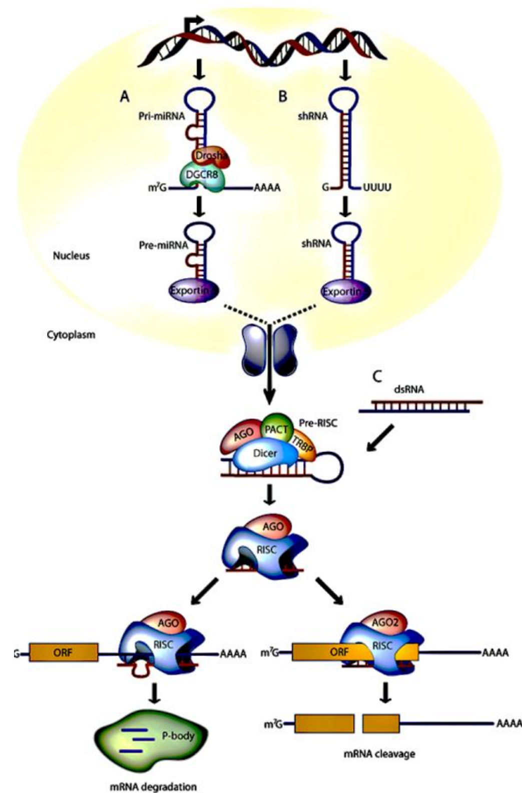

Figure 5 – miRNAs, shRNAs, and siRNAs pathways for RNAi in mammalian cells. (A) miRNA genes are transcribed by RNA polymerase II to generate the primary transcripts (pri-miRNAs) and processed into stem-loop precursor miRNAs of 70 nucleotides by the Drosha-DGCR8 complex, which are then exported by Exportin 5 to the cytoplasm. Upon export, Dicer participates in the second step of processing to produce miRNA duplexes of 22 nucleotides. The imperfectly complementary miRNA duplexes are associated to the AGO protein and are loaded into RISC, where the passenger strand is removed and the guide strand remains to target mRNA for silencing. The final products (RISC complex) act as guide molecules in translational control or cleavage of certain mRNAs. (B) Like miRNAs, engineered shRNAs are transcribed from DNA and undergo similar processing. However, the perfect Watson-Crick base-pairing between the guide strand and the target mRNA triggers AGO2-mediated cleavage of the mRNA target. shRNA expression cassette can be delivered by viral vectors such as retroviral vector, lentiviral vector, and adenoviral vector or it can be expressed in the nucleus from expression plasmids. (C) In contrast to miRNA and shRNAs, siRNAs are processed in the cytoplasm. But, all steps of siRNA and shRNA are the same after processing by Dicer/TRBP. siRNA can be artificially introduced into the cytoplasm in RNAi-based therapeutics either as a chemically synthesized siRNA duplex or Dicer-processed siRNA molecules. Viral and non-viral delivery systems are used for siRNA transfer into cells. (miRNAs, micro RNA; shRNAs, small hairpin RNA; siRNAs, small interfering RNA; RNAi, RNA interference; RISC, RNA-induced silencing complex; mRNA, messenger RNA). (Adapted from Burnett and Rossi (2012)).

The RNA interference technology is one of the most exciting biotechnology advances in the last decade. It has revolutionized biology research, including drug target discovery and revitalized interest in the clinical development of nucleic acid-based gene inhibition approaches. Theoretically, RNAi can silence the expression of mRNA for any gene, including growth factors, viral genes, or oncogenes and genes that were once considered therapeutically unreached by small molecule inhibitors. The great effectiveness and the simplicity of the design of a therapeutic siRNA, which only requires knowledge of the target gene's sequence, have contributed for the increased development and success of this RNA-based therapeutic approach. Moreover, the fact that siRNA-mediated RNAi mechanism takes place in the cytoplasm is a potential advantage over other gene regulation mechanisms that require penetrating the nucleus (Tokatlian and Segura, 2010). Therefore, several pharmaceutical companies are focusing on the development of RNAi-based therapeutics for the treatment of a wide range of diseases (Burnett and Rossi, 2012, Melnikova, 2007).

1.3.4. Aptamers

Many small RNAs can fold into three-dimensional structures that allow them to bind target proteins with high affinity and specificity. This additional feature of RNAs makes them tempting to consider as a therapeutic agent since they can bind to proteins and inhibit them in an analogous manner to protein antagonists. This idea of 'decoy' RNAs has been shown in preclinical work to slow down HIV replication, where *trans*-activation response region (TAR) and rev response element (RRE), two decoy RNAs, could be used to competitively inhibit viral protein function and replication. This suggested that other small-structured RNA molecules might be able to bind pathogenic target proteins and inhibit their activity (Kohn *et al.*, 1999). Meanwhile, synthetic short oligoribonucleotides sequences that bind to a specific target molecule with high affinity, called aptamers, emerged as potential molecules for both basic research and clinical purposes as macromolecular drugs. Aptamers are highly specific, relatively small in size, and non-immunogenic. Aptamers are essentially a chemical equivalent of antibodies. However, aptamers can be chemically synthesized to produce large quantities of these compounds for *in vivo* experimentation and clinical trials. Aptamers are usually created by iterative *in vitro* selection methods that isolate high-affinity RNA ligands from large pools of randomized RNA sequences (vast RNA shape libraries) that could bind to proteins and small molecules. The selection process was named SELEX (systematic evolution of ligands by exponential enrichment) (Ellington and Szostak, 1990), but natural aptamers also exist as part of a nucleic acid-based genetic regulatory element named riboswitch (Tucker and Breaker, 2005). Since the discovery of aptamers in the early 1990s, great efforts have been made to make them clinically relevant for diseases like cancer, HIV, and macular degeneration. For therapeutic purposes, aptamers can be used to bind and inhibit harmful molecules or serve as targeting ligands for nanomedical constructs. RNA aptamers, simplify the need for chemical conjugation or mixing with other moieties. Aptamers have been used as

ligands for specific delivery of siRNA to prostate cancer cells and lymphocytes (Ni *et al.*, 2011). The first aptamer based therapeutic was FDA approved in 2004 for the treatment of age-related macular degeneration and several other aptamers are currently being evaluated in clinical trials (Burnett and Rossi, 2012).

1.3.5. mRNA vaccination

Coding mRNA is emerging as a particularly attractive option in the development of new approaches for the treatment of cancer or infection diseases focusing on immunotherapies (Kreiter *et al.*, 2011). The use of genetic vaccination, in which a genetically engineered nucleic acid encoding an antigen is administered to an organism in order to stimulate an immunological response against the antigen (Tang *et al.*, 1992), is not a new therapeutic strategy. Vaccines based on nucleic acids are already known to stimulate all effectors of the adaptive immune response: B lymphocytes, cytotoxic T cells, and T helper cells have been exploited for the creation of prophylactic vaccines for infectious diseases and for cancer immunotherapy. However, vaccination approaches involving nucleic acids have focused mostly on DNA-based strategies due to the instability and rapid degradation of RNA in the human body, since RNA is prone to hydrolysis by ubiquitous ribonucleases. Nevertheless, researchers have been overcoming these limitations by developing strategies for stabilization and delivery of mRNA (Pascolo, 2008).

mRNA is a large family of RNA molecules that plays a fundamental and integral role in every living cell. mRNAs are carriers of genetic information from DNA to the ribosome, where they specify the amino acid sequence of the protein products of gene expression. Following transcription of mRNA by RNA polymerase, the mRNA is translated into a polymer of amino acids: a protein. Therefore, the concept of mRNA vaccination is to carry the information of an antigenic protein to be translated in the cell cytoplasm and therefore generate an immunological response. As an alternative to DNA-based vaccines, mRNA-based vaccines present additional safety features including no persistence, no integration in the genome and no autoimmune response. Moreover, mRNA which are generated by *in vitro* transcription, are easy to produce in large amounts and with very high purity. This feature facilitates the good manufacturing practices process and guaranties reproducibility (Pascolo, 2006).

Vaccination with mRNA can be achieved by several delivery methods, including direct injection of naked mRNA, injection of mRNA encapsulated in liposomes, gene gun delivery of mRNA loaded on gold beads, or *in vitro* transfection of the mRNA in cells followed by re-injection of the cells into the patients (Pascolo, 2008). This last strategy is generating a lot of interest in cancer immunotherapy. The approach uses tumour mRNA isolated from the patient by biopsy, amplified and combined with the immune cells from the patient’s bone marrow called dendritic cells (DCs). These cells internalize the mRNA and present the proteins encoded by the mRNA as an antigen on the cell surface, which ensures the stimulation of the

immune response to the tumour (figure 6). The preclinical experience suggests that vaccination with tumour RNA-transfected DCs may constitute a highly effective and broadly applicable treatment for patients with recurring cancer (Sullenger and Gilboa, 2002). Other technologies of direct injection of globin-stabilized mRNA are also being evaluated in human clinical trials (Pascolo, 2006).

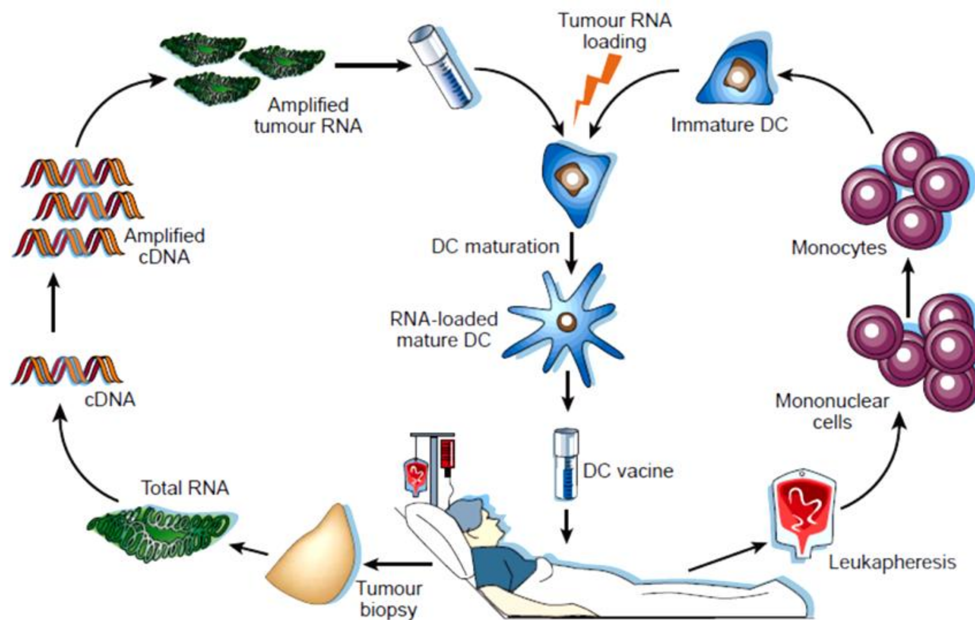

Figure 6 - Treatment of cancer patients with tumour RNA-transfected DCs. Tumour cells are removed from the patient and used to isolate, and if necessary amplify, tumour RNA. Blood cells are obtained from the patient and monocytes are isolated and their proliferation induced to immature DCs. Immature DCs are transfected with RNA and cultured to mature. The antigen-loaded DCs are then infused into the patient. (DCs, dendritic cells). (Adapted from Sullenger and Gilboa (2002)).

References

Brantl S. Antisense-RNA regulation and RNA interference. Biochim Biophys Acta. 2002; 1575:15-25.

Burnett JC, Rossi JJ. RNA-based therapeutics: current progress and future prospects. Chem Biol 2012; 19:60-71.

Dias N, Stein CA. Antisense oligonucleotides: basic concepts and mechanisms. Mol Cancer Ther. 2002; 1:347-55.

Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. Nature. 1990; 346:818-22.

Gascón AR, del Pozo-Rodríguez A, Solinís MA. Non-Viral Delivery Systems in Gene Therapy. In: Martin F, editor. Gene Therapy - Tools and Potential Applications: InTech; 2013.

Genzyme C. FDA News Release: FDA approves new orphan drug Kynamro to treat inherited cholesterol disorder. Food and Drug Administration; 2013.

Ginn SL, Alexander IE, Edelstein ML, Abedi MR, Wixon J. Gene therapy clinical trials worldwide to 2012 – an update. J Gene Med. 2013; 15:65-77.

Grillone LR, Lanz R. Fomivirsen. Drugs Today (Barc). 2001; 37:245-55.

Guo P, Coban O, Snead NM, Trebley J, Hoeprich S, Guo S, et al. Engineering RNA for targeted siRNA delivery and medical application. Adv Drug Deliv Rev. 2010; 62:650-66.

Hauswirth WW, Shaw LC, Whalen PO, Fritz JJ, White DA, Lewin AS. Inhibition of gene expression by ribozymes. Methods Mol Med. 2001; 47:105-24.

Kaji EH, Leiden JM. Gene and stem cell therapies. JAMA 2001; 285:545-50.

Khan AU. Ribozyme: A clinical tool. Clin Chim Acta. 2006; 367:20-7.

Kim VN. Small RNAs: classification, biogenesis, and function. Mol Cells. 2005; 19:1-15.

Kohn DB, Bauer G, Rice CR, Rothschild JC, Carbonaro DA, Valdez P, et al. A clinical trial of retroviral-mediated transfer of a rev-responsive element decoy gene into CD34(+) cells from the bone marrow of human immunodeficiency virus-1-infected children. Blood. 1999; 94:368-71.

Kreiter S, Diken M, Selmi A, Tureci O, Sahin U. Tumor vaccination using messenger RNA: prospects of a future therapy. Curr Opin Immunol. 2011; 23:399-406.

Lehman N. RNA in evolution. Wiley Interdiscip Rev RNA. 2010; 1:202-13.

Lin SL, Miller JD, Ying SY. Intronic microRNA (miRNA). J Biomed Biotechnol. 2006; 2006:26818.

Melnikova I. RNA-based therapies. Nat Rev Drug Discov. 2007; 6:863-4.

Ni X, Castanares M, Mukherjee A, Lupold SE. Nucleic acid aptamers: clinical applications and promising new horizons. Curr Med Chem. 2011; 18:4206-14.

O'Connor TP, Crystal RG. Genetic medicines: treatment strategies for hereditary disorders. Nat Rev Genet. 2006; 7:261-76.

Pascolo S. Vaccination with messenger RNA. Methods Mol Med. 2006; 127:23-40.

Pascolo S. Vaccination with Messenger RNA (mRNA). In: Bauer S, Hartmann G, editors. Toll-Like Receptors (TLRs) and Innate Immunity: Springer Berlin Heidelberg; 2008. p. 221-35.

Phylactou LA, Kilpatrick MW, Wood MJ. Ribozymes as therapeutic tools for genetic disease. Hum Mol Genet. 1998; 7:1649-53.

Sharp PA. The Centrality of RNA. Cell. 2009; 136:577-80.

Sheridan C. Gene therapy finds its niche. Nat Biotechnol. 2011; 29:121-8.

Soifer HS, Rossi JJ, Saetrom P. MicroRNAs in disease and potential therapeutic applications. Mol Ther. 2007; 15:2070-9.

Soll D, Nishimura S, Moore P. RNA. First ed. Oxford: Pergamon; 2001.

Strachan T, Read A. Gene therapy and other molecular genetic-based therapeutic approaches. Human Molecular Genetics 2nd edition ed. New York: Wiley-Liss; 1999

Sullenger BA, Gilboa E. Emerging clinical applications of RNA. Nature. 2002; 418:252-8.

Tang DC, DeVit M, Johnston SA. Genetic immunization is a simple method for eliciting an immune response. Nature. 1992; 356:152-4.

Tokatlian T, Segura T. siRNA applications in nanomedicine. Wiley Interdiscip Rev Nanomed Nanobiotechnol. 2010; 2:305-15.

Tucker BJ, Breaker RR. Riboswitches as versatile gene control elements. Curr Opin Struct Biol. 2005; 15:342-8.

Verma IM, Weitzman MD. Gene therapy: twenty-first century medicine. Annu Rev Biochem. 2005; 74:711-38.

Walther W, Stein U. Viral vectors for gene transfer: a review of their use in the treatment of human diseases. Drugs. 2000; 60:249-71.

Wang W, Li W, Ma N, Steinhoff G. Non-viral gene delivery methods. Curr Pharm Biotechnol. 2013; 14:46-60.

Wivel NA. Gene Therapy, Ethics, Somatic Cell Gene Therapy. Encyclopedia of Ethical, Legal and Policy Issues in Biotechnology John Wiley & Sons, Inc.; 2002.

Wolff JA, Lederberg J. An early history of gene transfer and therapy. Hum Gene Ther. 1994; 5:469-80.

Xiang S, Fruehauf J, Li CJ. Short hairpin RNA-expressing bacteria elicit RNA interference in mammals. Nat Biotechnol. 2006; 24:697-702.

2. Current issues in RNA preparation: approaching affinity chromatography into RNA purification challenges (Paper I)

R. Martins, J. A. Queiroz, F. Sousa

(Submitted for publication)

Current issues in RNA preparation: approaching affinity chromatography into RNA purification challenges

R. Martins, J. A. Queiroz, F. Sousa

CICS-UBI – Health Sciences Research Centre, University of Beira Interior, Av. Infante D. Henrique, 6200-506 Covilhã, Portugal.

Abstract

Research on RNA has led to many important biological discoveries and improvement of therapeutic technologies. From basic to applied research, many procedures employ pure and intact RNA molecules; however their isolation and purification are critical steps because of the easy degradability of RNA, which can impair chemical stability and biological functionality. The current techniques to isolate and purify RNA molecules still have several limitations and the requirement for new methods able to improve RNA quality to meet regulatory demands is growing. In fact, as basic research improves the understanding of biological roles of RNAs, biopharmaceutical industry starts to focus on it as a biotherapeutic tool.

Chromatographic bioseparation is the principal unit operation used for the purification of biological compounds and its application in biopharmaceutical manufacturing is well established. Thus, a number of chromatographic approaches have already been successfully developed for RNA purification. In particular and in view of the unequalled specificity, affinity chromatography has been recently applied, showing significant results and improvements in RNA purification processes.

Therefore, this paper discusses the importance and the progress of RNA isolation and purification procedures, considering the RNA applicability both in research and clinical field. Accordingly, recent investigations using affinity approaches based on the biorecognition between amino acids and RNA is focused, highlighting their potential contribution to overcome the challenges of RNA purification.

Keywords

Affinity chromatography, amino acids, isolation methods, purification, RNA.

1. Introduction

Until recently, RNA was overlooked compared to DNA or proteins, consigned to a simple intermediate role in the flow of information from genes to functioning molecules in living cells. RNA is now known to play many more functional roles and to be responsible for a multitude of essential biological processes (Sharp, 2009). In the last 20 years RNA was the subject of four Nobel prizes winning discoveries - 1989 for catalytic RNA, 1993 for splicing, 2006 for RNA interference (RNAi), and 2009 for ribosomal structure (Lehman, 2010) and new roles for RNA in biology continue to emerge at a glance. All of these discoveries have revealed so far that RNA is truly a remarkable and multi-talented cellular component with fundamental implication on biotic evolution and heredity. Furthermore, the widespread involvement of RNA in the regulation of numerous genes has highlighted its vast therapeutic potential (Burnett and Rossi, 2012). These and similar breakthroughs have led to the emergence of numerous types of RNA-based therapeutics either using RNA as a therapeutic agent or a therapeutic target. Table 1 shows potential therapeutic approaches for RNA, specifying the involved RNA or RNA-based molecules and their activity mechanism of activity.

The successful results of these novel therapeutic approaches are reinforcing the focus on RNA investigation and are rendering RNA molecules into new targets for pharmaceutical and biotechnological industries (Melnikova, 2007). Due to the increasing number of structural, biophysical and biomedical studies that require large quantities of homogeneous good-quality RNA, a widespread need to improve production scale and RNA isolation and purification schemes has been recognized.

As RNA emerges into the new class of biotherapeutic products, pharmaceutical-grade RNA produced under the current good manufacturing practice (cGMP) is crucial. Thus, it will be essential that the bioproduct fulfil the requirements of regulatory authorities such as Food and Drug Administration (FDA), European Medicine Agency (EMA) and World Health Organization (WHO). However, non-consensus still exists for regulation of mRNA vaccination and RNA oligonucleotides-based therapies. In the European Union, mRNA-based therapies are based on the regulation for advanced therapy medicinal products - EC No 1394/2007- (Guideline on Human Cell-based Medicinal Products) which refers to directive 2001/83/EC (Directive 2001/83/EC). Here, a ‘Gene therapy medicinal product is an active substance which contains or consists of a recombinant nucleic acid used in or administered to human beings’. In the United States of America, in contrast, mRNA vaccines are not categorized as gene therapy (Kreiter *et al.*, 2011). On the other hand, RNA oligonucleotides products, such as siRNAs or aptamers, are regulated as drugs under FDA’s Centre for Drug Evaluation and Research and are not considered as an advanced therapy in the European Union (not classified as gene therapy) (Srivatsa, 2012). Currently, none regulatory authority has formal guidelines available for RNA oligonucleotide products or mRNA molecules. The guidelines established for

human cell-based medicinal products (Guideline on Human Cell-based Medicinal Products) and DNA vaccines (Guidance for Industry: Consideration for plasmid DNA vaccines for infection disease indications) are providing the guidance in the regulatory framework for RNA-based therapies (Srivatsa, 2012, Weide *et al.*, 2008). Nevertheless, those guidelines do not focus in particular features of RNA, such that as shown in clinical trials, RNA-based therapies do not confer the risks of integration into the genome. The scientific community, in the form of volunteer members from the industry and regulatory agencies worldwide, are actively engaged in addressing topics such as quality specifications and impurities in RNA bioproducts (Srivatsa, 2011).

Consequently, economically feasible processes for RNA isolation and purification, as well as the implementation of methodologies able to control RNA quality suitable for industrial manufacturing, will be increasingly necessary, especially when the RNA products are finally released to the market. However, the isolation and purification of RNA are critical steps because of the easy degradability of RNA, consequence of the peculiar structural chemistry, which can impair chemical stability and biological functionality, and can limit the success of subsequent RNA investigations.

Therefore, this paper focuses the challenging task of isolating and purifying RNA molecules. The current state of the isolation and purification methodologies used for RNA preparation will be discussed regarding the growing demands in RNA applicability. Moreover, promising affinity approaches based on chromatographic purification exploiting the biorecognition between amino acids ligands and RNA molecules will be introduced. These new strategies bring new insights into the way RNA can be accurately purified, contributing to the future development of new and more robust bioseparation methods.

Table 1 - RNA molecules with therapeutic involvement. (Refs., References).

RNA type	Cell function	Therapeutic concept	Refs.
		Protein synthesis	
Messenger RNA (mRNA)	Codes for protein	Vaccination	Kreiter *et al.* (2011)
Ribossomal RNA (rRNA)	Translation	Antibiotic target	Tenson and Mankin (2006)
Transfer RNA (tRNA)	Translation Apoptosis regulation	Understand many human diseases	Belostotsky *et al.* (2012), Raam and Salvesen (2010)
		Post-transcriptional modification or DNA replication	
Small nuclear RNA (snRNA)	Splicing and other functions	Understand many human diseases	Matera *et al.* (2007)
Small nucleolar RNA (snoRNA)	Nucleotide modification of RNAs	Understand many human diseases	Kiss (2001)
		Regulation	
Antisense RNA	Transcriptional attenuation mRNA stabilization or degradation Translation block	Inhibitor of mRNA translation	Brantl (2002), Dias and Stein (2002)
MicroRNA (miRNA)	mRNA cleavage and Translation repression	Gene silencing	Kusenda *et al.* (2006) Lin *et al.* (2006)
Small interfering RNA (siRNA)	mRNA cleavage and Translation repression	Gene silencing	Doench *et al.* (2003), Ghildiyal and Zamore (2009)
Ribozymes	RNA enzyme. Catalyze RNA cleavage and ligation reactions	mRNA reprogramming and repair	Phylactou *et al.* (1998)
Riboswitch	Regulate gene expression by binding to small metabolites	Regulate gene expression Antibacterial drug target	Blount and Breaker (2006), Tucker and Breaker (2005), Wittmann and Suess (2012)
Aptamers	Oligoribonucleotide part of a riboswitch that binds to a specific target molecule with high affinity	Decoy mechanism that inhibits various target proteins	Mayer (2009), Ni *et al.* (2011)

2. RNA chemistry

RNA has a number of unique chemical characteristics that have profound structural consequences with remarkable implications in cell biology and are a real challenge in research activities.

RNA is a polymer organized in a long chain of ribonucleotides monophosphates, but it resembles DNA in many ways (Figure 1). RNA shares the same chemical units as DNA, except that RNA has the nucleotide uracil (U) instead of thymine (T) and the sugar molecule is a ribose rather than a desoxyribose. RNA has a hydroxyl group at the 2′ position of the ribose while DNA lacks that group. This difference is manifested in the dissimilar shape that these molecules adopt when they are base-paired into a double helix. RNA takes on the geometry structure referred as an A - form helix while DNA takes on the B - form. But, unlike double-strand DNA, RNA often comes as only a single-strand and is quite flexible (Murray *et al.*, 2000). RNA can twist itself into a variety of complex structures. Its propensity to form secondary structures facilitates RNA interactions with other molecules by covering some sequences and exposing others for recognition. Besides, RNA can assume tertiary structures that present surfaces for interactions and contain internal environments that create binding sites for metal ions, so that they can promote catalytic reactions (Soll *et al.*, 2001).

The high chemical reactivity of RNA provides more instability to the molecule, increasing the susceptibility to degradation. This instability is very important for cells, as they can change their patterns of protein synthesis very quickly in response to biological needs (Arraiano *et al.*, 2010).

These properties consign RNA versatility in cellular processes, namely in gene regulation, which open the possibility of exploring new therapeutic opportunities (Sullenger and Gilboa, 2002). On the other hand, the peculiar three-dimensional compaction and structural instability of RNA are huge challenges in laboratory, as the biological activity and integrity can be easily compromised during extraction and purification procedures. Thus, improved methodologies for recovering RNA with high quality is a constant concern (Easton *et al.*, 2010, Fleige and Pfaffl, 2006, Jahn *et al.*, 2008).

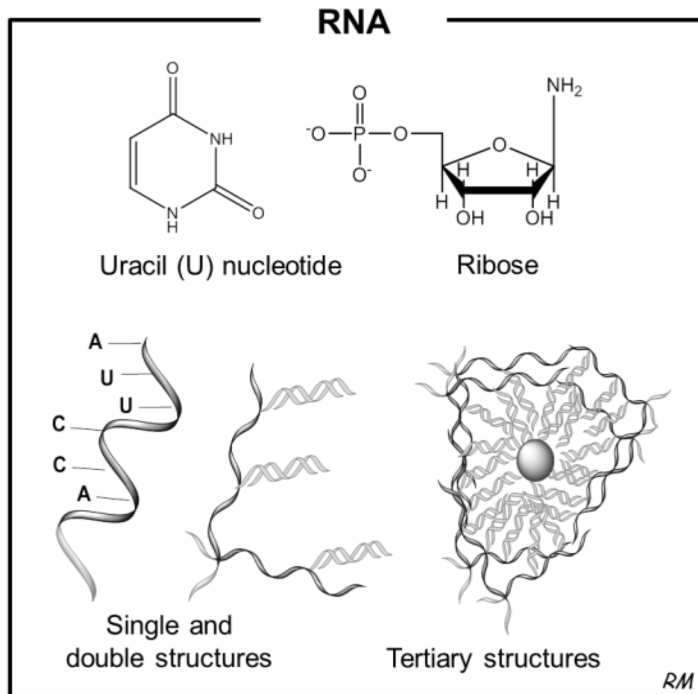

Figure 1- Structural characteristic of RNA. RNA differs from DNA in the nucleotide uracil, as a thymine exists in DNA, as well as the constituent sugar molecule that is a ribose in RNA and a desoxyribose in DNA. In addition, three-dimensional structures of RNA also greatly differ from that of DNA. While DNA molecules take a double helix structure, which is formed through hydrogen bonding between the bases of two DNA strands, RNA molecules, which are originally synthesized as single-strands, exhibit different conformational structures where single and double-strands coexist in one RNA molecule because they often partially fold themselves to form double-strands. Compact tertiary conformations can exhibit internal surfaces for interaction with other molecules.

3. RNA isolation and purification

The development of methodologies for the isolation and characterization of RNA has been compared to the methods used for the analysis of its macromolecular counterparts, DNA and proteins (Lehman, 2010). The extraction, isolation, and analysis of RNA are routinely more difficult in comparison to that required for DNA. As raised before, RNA chemistry add complexity to sample preparation because the ubiquitous presence of RNA-degrading enzymes (RNases) both in biological samples and in laboratory environment easily degrade RNA, compromising the integrity and biological activity of RNA molecules (Jahn *et al.*, 2008). Therefore, minimizing RNA degradation by protecting it against RNases requires that all glassware, plastic ware, instrument tubing and reagents be RNase free. Additionally, among the many challenges is the need for maximizing recovery yield, while removing unwanted components, minimizing sample transfers, and avoiding non-specific binding to containers (McGinnis *et al.*, 2012). Therefore, the quality and quantity of RNA preparations are the main concerns of isolation procedures, since the lack of integrity, the presence of contaminants or the low RNA quantity may strongly compromise the success of several RNA based-procedures in basic and clinical research (Bustin and Nolan, 2004). Additional challenges also emerge with the advance of clinical trials using RNA intended to be administered in humans.

Presently, RNA molecules can be obtained by extraction from a biological matrix, such as cells or tissues, or they can be produced by chemical or enzymatic (*in vitro* transcription) synthesis. Chemical synthesis is normally used for the generation of short oligoribonucleotides (<50 nucleotides) while *in vitro* transcription can produce longer RNAs. Synthetized RNAs are being greatly employed in structural, biochemical and biophysical studies (Doudna, 1997, Scott and Hennig, 2008) as well as in the development of new therapeutic approaches by RNA interfering technology, RNA aptamers, ribozymes or mRNA vaccination (Lorenzi *et al.*, 2010, Wittmann and Suess, 2012, Zeng *et al.*, 2002). In this case, the final RNA product needs to be purified from impurities derived from the synthesis process. These impurities are, besides enzymes, nucleotides, salts or buffer, aberrant oligonucleotides. Longer oligoribonucleotides are more contaminated with aberrant species than short ones. The failure products are prematurely halted as shorter oligonucleotides. Some are mismatch failure sequences where there is missing nucleotides in the middle of sequence, rather than at the end. Other by-products of synthesis may have greater molecular weight than the target oligoribonucleotide (heterogeneous RNAs in length). This is a result of incomplete post-synthesis deprotection, or due to the branching of an oligonucleotide backbone during the synthesis (Sherlin *et al.*, 2001).

On the other hand, biological RNAs are preferably used in basic research for the study of cellular mechanisms, as they reflect intrinsic cell´s features (Dunin-Horkawicz *et al.*, 2006), in clinical investigations for pharmacokinetic and pharmacodynamics analysis (Yu *et al.*, 2001)

and in some strategies of mRNA vaccination using bulk tumour mRNA (Kyte *et al.*, 2006). In RNA extraction from a biological matrix, the main principle is the disruption of cells and subsequent elimination of host contaminants, such as genomic DNA (gDNA) and proteins, in order to obtain intact and pure RNA molecules (Chomczynski and Sacchi, 2006).

Many methods have been developed in an attempt to circumvent the several challenges in purifying RNA molecules and to achieve the goal of good-quality RNA (Farrell, 2012, McGookin, 1988, Vomelova *et al.*, 2009, Warren and Vella, 1995). Therefore, considering the recent developments in RNA understanding and the growing demand on its purification, the next discussion intends to briefly introduce the isolation and purification methods used in RNA preparation. This description will draw particular attention to the main problems that can compromise the success of RNA research or that can make the procedures not suitable to obtain RNA to be further applied in clinical investigation, also focusing on the cost-effectiveness for preparative-scale or large-scale industrial applications (Table 2 and 3).

3.1. RNA isolation methods

Traditionally two types of isolations are used in RNA preparations, (1) chemical extraction using denaturing agents and organic solvent precipitation and (2) solid-phase extraction by immobilizing RNA on a glass support. These methods generally include a chaotropic agent, denaturant, or other chemical in the lysis step to inactivate RNases.

Chemical extraction that involves acid phenol/chloroform extraction is the most currently employed method either performed with home-made solutions or commercial ready-to-use reagents, because it leads to high recovery yields and purity of total RNA (RNAt) (Chomczynski and Sacchi, 2006). However, these extractions are extremely toxic and hazardous and highly operator dependent. Sometimes they can involve up to six or more steps and three sample transfers leading to time consuming and laborious RNA preparations. Although this method is almost always included in RNA purification schemes, the organic solvents such as alcohols and phenol/chloroform may interfere in the majority of routine molecular biology techniques because they can inhibit enzymes activity. Moreover, as those compounds convey health risks, this procedure is not definitely tolerable for the welfare of the researcher and should not be an integrant part of the process for a therapeutic formulation. However, RNAs that are being employed in clinical trials are often extracted using phenol/chloroform (Kyte *et al.*, 2006). This should, certainly, be considered and advised by regulatory authorities.

The other methods for RNA isolation are based in solid-phase purification using silica membranes, as prefilled columns or as magnetic beads, in combination with phenol/guanidine-based lysis of samples. This technology was developed to offer safer and simpler operations, as they are amenable to automation using liquid handling robotics (Wen

et al., 2008). In fact, these techniques significantly reduce sample preparation time, but can still involve multiple sample transfers and time-consuming evaporation steps. Low RNA yields are often obtained due to the low binding capacity of the cartridges, which can be easily overloaded (Chatterjee *et al.*, 2010, Wen *et al.*, 2008). Moreover, these processes lack specificity as they are non-discriminatory for DNA or RNA, and in most cases the isolation of pure RNA is achieved by a secondary enrichment, either through enzymatic removal of DNA (DNase treatment) or by a second step using specific columns (Chatterjee *et al.*, 2010). These limitations can greatly diminish the success of several molecular biology investigations, in particular gene expression analysis, because DNA contamination can interfere with real time-PCR techniques compromising the interpretation of results (Laurell *et al.*, 2012). In addition, the treatment with enzymes, such as commercial DNases that are often animal-derived, is not prudent in therapeutic applications (Guidance for Industry: Consideration for plasmid DNA vaccines for infection disease indications).

Recent developments in high throughput technologies led to the design of large and complex instrumentation using silica technology to simplify the isolation of nucleic acids. Automation for nucleic acids extraction offers high-throughput sample processing, high quality and effective purification since the human impact is constricted to a minimum (Chatterjee *et al.*, 2010). Therefore, less contaminants or inequalities concerning time, volumes of chemicals or improper handling are avoided. Automation therefore increases the efficiency, accuracy and velocity of RNA purification (Knepp *et al.*, 2003). Nevertheless, cross contaminations can occur by the sample collector arm. In these systems, the principles of purification are based in solid-phase extraction, which presents some limitations, as previously mentioned. While the high cost disadvantage of these systems could be overcome by the high-throughput of automation, the lack of efficacy in the isolation method can commit the investment. Actually, automated extraction systems may even be advantageous to some laboratories, since consistent and reproducible purified samples in sufficient quality and purity are achieved for subsequent application (Loens *et al.*, 2007). However, in a biotechnological process, these systems will not be cost-effective in attaining RNA-based therapeutic formulations.

The use of preparative denaturing polyacrylamide gel electrophoresis (PAGE) has also been extensively described to complement the schemes described above in the preparation of RNA, especially synthetic RNA (Doudna, 1997, Hagen and Young, 1974, Price *et al.*, 1995, Sherlin *et al.*, 2001). Preparative denaturing PAGE has been successful and continues to be the most popular method to resolve RNA samples, but it presents several disadvantages. The protocol is very lengthy as after denaturing PAGE, the desired RNA must be eluted from a gel matrix, concentrated, equilibrated in the desired buffer, and refolded (Doudna, 1997). Moreover, this introduces acrylamide contaminants, which are difficult to remove from RNA products and can interfere with RNA analysis, namely by reducing the information content on nuclear

magnetic resonance (NMR). Although highly purified, the RNAs achieved by PAGE purification show low yields since RNA can irreversibly aggregate due to many precipitation steps, resulting in incomplete denaturation during electrophoresis and low accumulation in the desired gel band (Kim *et al.*, 2007, Lukavsky and Puglisi, 2004). Moreover, denaturation of RNA molecules is another issue for RNA applicability. For many RNAs, significant time is spent optimizing refolding conditions to minimize unproductive conformations. Some well-known RNAs, such as *Escherichia coli* (*E. coli*) $tRNA^{Phe}$, cannot be refolded into a conformational homogeneous and active population (Kieft and Batey, 2004). Therefore, native purification techniques should be favoured. From the therapeutic viewpoint, PAGE purification is rather impracticable for many reasons. First, the prominent possibility of acrylamide contaminants in RNA preparations is not adequate or safe for human therapy. Furthermore, structural modifications due to denaturation can greatly compromise the effectiveness of RNA therapeutic action.

Precipitation with high concentration of lithium chloride (LiCl) is another methodology often used for synthetic RNA isolation (Baker *et al.*, 1997, Nilsen, 2012, Pascolo, 2008). This method can be used in the separation of small RNAs (<100 nucleotides long) (i.e., siRNA and miRNAs) from other long RNAs or RNA-based impurities. Small RNAs are enriched in the supernatant phase, while remaining impurities or all other long RNAs tend to precipitate (Baker *et al.*, 1997, Nilsen, 2012). However, the efficient fractionation typically requires repetition of the LiCl precipitation procedure several times and the quality of resulting preparation is in general dependent on the secondary or tertiary conformation of RNA molecules (Romanovskaya *et al.*, 2013). Other applications describe the use of LiCl precipitation in *in vitro* transcripts reaction to eliminate most nucleotides, cap, deoxynucleotides, oligodeoxynucleotides and proteins by selectively precipitating RNA of more than approximately 100 bases, as mRNA transcripts (Pascolo, 2008). Nonetheless, most of the time, this precipitation does not completely remove trace of contaminating plasmid DNA (pDNA) or proteins and worse it introduces lithium metal into preparations. Thus, to ensure long RNA purity, phenol/chloroform extraction often accompanies LiCl precipitation. Although this method can be suitable for preparation of mRNA molecules for research, it is not recommended to be used in clinical applications, as in vaccination approaches, since the safeness of those preparations are not guaranteed to be applied in humans (Pascolo, 2008).

Table 2 - Isolation techniques used in RNA preparation. (Refs., References; RNAses, Ribonucleases; DNase, Desoxyribonuclease; nt, nucleotides; PAGE, polyacrylamide gel electrophoresis; cGMP, current good manufacture practice; LiCl, lithium chloride).

Method	Principle	Advantages	Disadvantages	Refs.
Phenol/chloroform extraction	Chaotrope helps cells lysis and inactivates RNases. Proteins are denatured and removed in the organic phase during phenol /chloroform extraction. RNA is precipitated with alcohol and salt.	Inexpensive Enhanced protection against RNases Good yields and purity	Requirement of toxic chemicals Time-consuming Highly operator dependent Inhibit enzyme activity	Chomczynski and Sacchi (2006)
Solid phase extraction	Chaotrope helps cells lysis and inactivates RNases. RNA and DNA bind to silica in the presence of chaotrope. DNA is digested with DNase. Contaminants are washed away and RNA is eluted with low ionic strength	Fast Does not require use of organic solvents or alcohol precipitation Amenable to automation	Low binding capacities Does not discriminate between RNA or DNA Requires DNase treatment	Wen *et al.* (2008)
Preparative denaturing PAGE	Separates molecules based on their electrical charge and hydrodynamic properties, which are a function of chain length. The desired RNA is eluted from the gel matrix, concentrated, equilibrated in buffer and refolded.	Highly purified product	Time-consuming Introduces contaminants Uses denaturing conditions Low yields	Doudna (1997), Hagen and Young (1974)
Lithium chloride precipitation	Selective separation of long or short RNA sequences from impurities. Elevated concentration of LiCl is added to impure RNA preparations follow by incubation at -20°C for several hours or overnight. Precipitates or supernatants are recovered according to required RNA type.	Separation of small RNAs from long RNAs Recovery of long RNAs from impurities of *in vitro* transcription synthesis	Inefficient precipitations Introduces lithium metal into preparations Employ phenol/chloroform extraction to improve isolation Time-consuming	Baker *et al.* (1997), Nilsen (2012), Pascolo (2008), Romanovskaya *et al.* (2013)

3.2. RNA purification methods

In recent years, several strategies employing chromatographic techniques have been explored to overcome the increasing challenges in RNA purification. Advances in synthetic RNA chemistry brought more efficient and affordable methods for large scale production of RNA. Accordingly, synthetic techniques opened the possibility of inducing chemical modifications on RNA molecules, reducing RNA instability and degradation (El-Sagheer and Brown, 2010), benefiting structural, biochemical and biophysical studies, in techniques as crystallography, spectroscopy, calorimetry or NMR, and many therapeutic investigations that employ regulatory RNAs, as siRNA, miRNA, ribozymes and aptamers. However, not so many efforts have been made for the improvement of biological RNAs purification.

Reversed-phase (RP) and ion-pairing (IP) RP (El-Sagheer and Brown, 2010, Maina *et al.*, 2013, Maroney *et al.*, 2007), as well as anion-exchange chromatography (AEC) have been extensively described for RNA purification (reviewed in (McGinnis *et al.*, 2012)). In addition, size exclusion chromatography (SEC) has also accompanied the development of schemes for RNA preparation (McKenna *et al.*, 2007).

RP chromatography relies on hydrophobicity as a mechanism of separation while IP chromatography describes a RP-based improved technique in which an ion-pairing reagent, such as a tri- or tetraalkylammonium salt, is capable of associating with the nonpolar stationary phase through dynamic hydrophobic interactions, allowing enhanced resolution (Cramer *et al.*, 2011). In fact, a number of studies for the purification and analysis of RNA is already available, demonstrating the versatility of IP RP HPLC in different applications with synthetic oligoribonucleotides or biological RNA (reviewed in Dickman (2011b)). Although these analytical studies include reliable information on RNA purification, very few of these methods have been scaled up to a preparative context. The reason for this is that resins have only moderate loading capacity and the cost of mobile-phase components and separation matrices are considerably high. In addition, some organic solvents and ion-pairing agents are difficult to remove from the purified oligonucleotides after chromatography and may even require further chromatographic purification (Gilar, 2001). Furthermore, to avoid aggregation of self-complementary or GC-rich oligonucleotides, RP HPLC columns are often placed in a column oven at 60°C for temporarily destroying secondary structures to improve resolution (Cramer *et al.*, 2011), which can be disadvantageous to maintain the target RNA stability and biological activity.

Therefore, recent investigations have explored the possibility to develop more rigorous and complex IP RP HPLC methods, since great resolving properties of RP columns were evidenced in analytical RNA separation. A recent approach describes the combination of SEC (porous silica with a hydrophilic polymeric coating) and IP RP HPLC (polystyrene-divinyl benzene beads) in a robust multidimensional platform to resolve, isolate and quantify non-coding RNA

species (ncRNA) in cells or tissues samples from several sources (Chionh *et al.*, 2013). The method takes advantage of the strengths of two types of HPLC techniques, thus increasing the resolution of ncRNA across a wide size range. Many technical issues concerning ncRNA preparations were improved, including the analysis speed, the application of non-denaturing conditions, easy quantification, sensitivity, easy automation for fraction collection and use of simple solvent mobile phases that do not contaminate the sample.

In addition, another recent strategy employed IP RP HPLC technique for the generation of a therapeutic mRNA pursuing vaccination (Karikó *et al.*, 2011), which purify long *in vitro*-transcribed mRNA at a preparative scale. Despite the use of some organic solvents in the procedure, the volatility of the ion-pairing agent was higher to be easily removed from samples through evaporation and non-denaturing conditions were employed. The successful applicability of the IP RP HPLC purification technique was characterized by the immunogenicity of the samples, which demonstrated the effectiveness of the purified RNA at an immunological level. This strategy can conveys important information about the implementation of future purification strategies suitable for clinical applications (Kreiter *et al.*, 2011).

Regardless the success of these chromatographic techniques, the requirement of toxic solvents continues to be their major weakness. The environmental impact and cost that the use of organic compounds entails can be highly inconvenient for a biotechnological industry besides, recovered RNA products might require further treatment to be adequate for therapeutic purposes.

SEC has also accompanied the advances in RNA purification. This technique is able to distinguish the components of a mixture of RNA on the basis of their molecular size, and is the simplest form of chromatography for oligoribonucleotides purification. SEC has contributed as a polishing step in removing salts from short oligoribonucleotides that have been purified by HPLC. This extra form of desalting prevents cytotoxic effects from trace synthesis by-products or trace solvents which may carry over from purification (Booy *et al.*, 2012). In addition, this technique has been explored as an alternative method to preparative denaturing PAGE in purifying homogeneous-length RNA obtained by *in vitro* transcription (Kim *et al.*, 2007, Lukavsky and Puglisi, 2004, McKenna *et al.*, 2007). Puglisi and workers developed SEC-based purification schemes using fast performance liquid chromatography (FPLC) systems that allowed the efficient elimination of unreacted nucleotides, enzymes, short abortive transcripts, and the high molecular weight pDNA template from the desired RNA product. The approaches were performed under non-denaturing conditions, which allows the exclusive separation of monomeric RNA from oligomerized RNA, and avoided harsh precipitation steps that may cause RNA aggregation and degradation (McKenna *et al.*, 2007). However, SEC-based methods still require several time-consuming preparatory steps, such as phenol/chloroform

extractions to remove proteins followed by desalting and sample concentration (Easton *et al.*, 2010).

Chromatographic studies on anion exchange matrices have been significantly explored because of the polyanionic nature of RNA molecules. AEC mechanism relies primarily on reversible electrostatic interactions that can be manipulated using an oppositely charged buffer, typically simple salts and buffers are needed for separation. Furthermore, anion exchange matrices are also described to have some inherent hydrophobicity which favours oligonucleotides separation. AEC selectivity for oligonucleotides was found to be based on molecular size and sequence and the success of separating double and single-stranded nucleic acids relies on the distinct hydrophobic behaviours of the G-C, A-T and A-U base pairs (Gjerde *et al.*, 2009). Additionally, AEC demonstrates simplicity on the operations compared to RP, since the separation is performed in aqueous conditions without the use of high cost eluents, at low to moderate operating pressures and the secondary structures that interfere with resolution are avoided using mild conditions to perform the purification. Hence, oligonucleotides purification during mid- or large-scale manufacturing is typically performed using AEC rather than RP chromatography (Shanagar, 2005).

Nonetheless, in some AEC strategies the necessity of denaturing conditions either by the use of harsh alkaline conditions or elevated temperatures to accomplish full resolution of oligoribonucleotides persists (Thayer *et al.*, 2011a). As already mentioned, denaturation often leads to misfolded or aggregated material or even leads to RNA degradation, which is unsuitable for further usage. Therefore, the methods can reveal not cost-effective in the preparation of RNA. Example of this is the preparation of double-strand RNAs (dsRNA), such as siRNA or miRNA. The chromatographic purification is generally performed for the single-strand intermediates rather than the final duplex, under denaturing conditions and therefore further purification of the final siRNA duplex in native form is required and is determined by SEC (Ge *et al.*, 2003, Thayer *et al.*, 2011a). Despite the successful application of this strategy in the manufacture of virtually all therapeutic duplex oligoribonucleotides (Srivatsa, 2011), the workload and process time to purify the separated single-strands, rather than the final duplex is increased, which also increases process economics.

New developments in AEC purification strategies intended to address large scale duplex preparation under native conditions are required. Good separations had already been achieved in analytical non-denaturing AEC, and recently preparative methods were also described (Noll *et al.*, 2011). Therefore, the scale up production of dsRNA has been optimized in two commercial AEC resins using the same quaternary amine as functional group, but differing on matrix chemistry: one material consisted on a polystyrene/divinyl benzene beads and the other on poly-(methyl-methacrylate). Both columns allowed the removal of most non-hybridized strands as well as non-optimal duplexes from the hybridized duplex siRNA, with higher yields than in parallel assays to purify single-strands intermediates before annealing.

Thus, subsequent duplex purification enabled the reduction of chromatographic steps, as the final SEC is no longer required, and annealing was significantly simplified.

Furthermore, convective interaction media (CIM) monolithic columns immobilized with strong anion exchange quaternary amines were also employed in the development of improved dsRNA purification methods under non-denaturing conditions (Romanovskaya *et al.*, 2013). Differently from chemical synthesis, this work describes the purification of siRNAs from a synthetic pool of functional siRNA produced by the enzymatic generation of long dsRNA molecules followed by digestion with a recombinant Dicer enzyme *in vitro*. Multiple contaminants, such as non-hybridized or unprocessed single-strand RNA, undigested or abortive dsRNA molecules, nucleoside triphosphates, that are obtained in transcription reaction were removed and siRNA was recovered at superior speed and scalability compared to the conventional bead-based AEC columns and traditional stepwise LiCl precipitation. Therefore, CIM monolithic columns were found to be suitable for industrial scale siRNA and dsRNA purification, while retaining fast processing times and high resolution.

With the improvement of synthetic RNAs by chemical modification, other challenges in RNA purification arose. Apart the regular impurities, many isomerizations, including 2′,5′-linkages, phosphoramidate-linked RNA, phosphorothioate-linked, 2′-cyanoethoxymethyl-protected RNA and oligonucleotides harboring phosphorothioate diastereoisomers, may also be formed during chemical synthesis (Thayer *et al.*, 2011b). Pellicular anion-exchangers (pAE) have been reported in literature for its ability to analytically resolve those isomers. pAE phases were designed to place the ion exchange sites only on the bead surface in order to overcome the mass transfer limitation of porous beads. While they were capable of very high chromatographic efficiency, they exhibit very low capacity, and they may harbor relatively high nonspecific interactions (Stillian and Pohl, 1990). Recently, a surface-functionalized monolith coated with pAE nanobeads (latexes) was developed, which intended to enhance purifications of single and dsRNA, aptamers, and nucleic acids that may harbor isobaric linkage isomers, such as phosphorothioate diastereoisomers and similar isobaric oligonucleotide variants (Thayer *et al.*, 2010). These new resins combine the selectivity and mass transfer characteristics of pAE with the preparative capacity of monoliths in order to use the successful analytical separations of pAE at a lab-scale. With this strategy it could be possible to evaluate and characterize the oligonucleotides regarding its safety and effectiveness as therapeutic and diagnostic products. However, further efforts are need to scale up these methods so that the resultant RNA preparations can be readily used as therapeutic products.

AEC has also been used as an alternative methodology to purify large scale *in vitro* transcribed long RNAs in order to circumvent the laborious and hazardous denaturing gel electrophoresis. Commonly weak and strong anion exchangers were reported for this purpose. One method describes a series of three weak anion-exchange diethylaminoethyl (DEAE)-

Sepharose FPLC columns that can purify RNA from the direct loaded of crude transcription reactions under non-denaturing conditions using mobile phases of sodium phosphate and sodium chloride in concentration up to 1M (Easton *et al.*, 2010). This technique enables to remove free nucleotides, short abortive transcripts, linearized plasmids, and enzymes from the desirable transcripts that can range in length from 30 to 500 nucleotides. Actually, this method does not purify RNAs with homogeneous 3'ends often crucial for X-ray crystallographic studies, but it yields structurally homogeneous, natively folded RNA, which is essential for many biochemical and biophysical applications.

Improving upon this strategy, a strong exchanger of quaternary amines was described as a useful tool for transcripts purification (Koubek *et al.*, 2013). The method was successfully applied to a variety of RNAs with different chain lengths, including $tRNA^{Cys}$, ribozyme transcripts, and 4.5S RNA transcripts. All protocols were performed with only slight variations without the need for special buffer systems, demonstrating the robustness of strong anion-exchange support. When compared to DEAE Sepharose column, the resolving power of the strong AEC was higher for the separation of RNA transcripts from the short oligonucleotide templates. In addition, the purification by strong AEC led to the enrichment of the functional tRNA from run-off transcripts. Nevertheless, the use of higher salt elution interfered with the RNA folding.

Table 3 - Purification methods used in RNA preparation. (Refs., References; nt, nucleotides; ncRNA, non-coding RNA; PAGE, polyacrylamide gel electrophoresis; SEC, size exclusion chromatography; dsRNA, double-stranded RNA; CIM, convective interaction media; AE, anion-exchange; SPE, solid phase extraction).

Chromatographic Method	Matrix	Advantages	Disadvantages	Refs.
Reversed-phase	Modified silica with hydrocarbon chains (normally C8 or C18)	Based on differences in hydrophobicity High resolution Efficient purification	Limited to short RNA sequences (<50nt) Requirement of toxic chemicals Uses denaturing conditions	McGinnis *et al.* (2012)

Ion-pairing reversed-phase	Polystyrene-divinyl benzene beads	In combination with SEC provide more accurate purification of biological ncRNAs Synthetic mRNA purification at preparative scale Non-denaturing conditions Use of simpler solvents	Difficult to scale up Still use toxic solvents	Chionh *et al.* (2013), Dickman (2011a), Karikó *et al.* (2011)
Size exclusion	Cross-linked dextran Cross-linked agarose gel	Powerful as final polishing step in oligoribonucleotides purified by HPLC Effective alternative to preparative PAGE purification of *in vitro* transcribed RNA Non-denaturing conditions	Time-consuming Highly operator dependent	Kim *et al.* (2007), Lukavsky and Puglisi (2004), McKenna *et al.* (2007)
Anion-exchange	Strong quaternary amine	Purification of synthetic dsRNA, rather than separated single-strands Effective alternative to preparative PAGE purification of *in vitro* transcribed RNA Non-denaturing conditions Large scale preparation of dsRNA using CIM monoliths columns	RNA folding can be compromised by chromatographic conditions	Noll *et al.* (2011) Koubek *et al.* (2013) Romanovskaya *et al.* (2013)
	Weak DEAE	Effective alternative to preparative PAGE purification of *in vitro* transcribed RNA Natively folded RNA Non-denaturing conditions Scaled up methods	Three columns in series to achieve purification Does not purify RNAs with homogeneous 3'ends	Easton *et al.* (2010)
	Pelicular AE (latex)	Separates short synthetic RNAs from isomers impurities Evaluation and characterization of therapeutic synthetic RNAs High efficiencies and capacities Lab-scale	Difficult to scale up	Thayer *et al.* (2010)

4. RNA affinity chromatography

The use of affinity chromatography (AC) has become invaluable as a platform in the development of therapeutically useful products because of the unique property to simulate and exploit natural biological processes such as molecular recognition for the selective purification of the target molecule (Lowe *et al.*, 2001). Actually, in general, affinity techniques and concepts have become essential tools in bioscience with applications spanning from protein purification, protein interaction mapping, development of biopharmaceuticals and diagnostics, as well as new developments in genomics and proteomics (Uhlen, 2008).

The major advantage of AC relies in its distinctive capacity of establishing a combination of different types of interactions, including electrostatic or hydrophobic interactions, van der Waals forces, or hydrogen bonding, due to the specific biorecognition of the target biomolecule by the specific ligand, which gives the technique an extremely high selectivity and high resolution. Thus, the target biomolecule can easily be separated from a crude sample in one step with thousands fold of increase in purity and high recovery (Roque and Lowe, 2008). The choice of matrix and conditions for purification will depend on the molecular properties of biomolecules and the physiochemical and thermodynamic nature of their molecular interactions. Elution steps can be performed either specifically, using a competitive ligand, or non-specifically, by changing the pH, ionic strength or polarity depending on the matrix used and the chemical characteristics of the biomolecules (Magdeldin and MoserTech, 2012). Overall, affinity methods have the advantages of eliminating additional steps, increasing yields and improve process economics.

One trend that has always been present in AC has been the search for more selective, robust, and/or reproducible ligands. Although the design of selective ligands for the purification of biomolecules is complex, time consuming and expensive, their implementation into AC processes would result in important economic advantages (Labrou, 2003), such as the reduction of downstream steps and the improvement of the product quality, therefore justifying the initial investments.

These are the main reasons justifying that AC processes have accompanied the growing demands for highly selective, reliable and economical processes to conform to strict quality assurance in production of therapeutic biomolecules (Ayyar *et al.*, 2012, Sousa *et al.*, 2008b). Therefore, and in view of RNA peculiar chemical structure and versatility in biological function that includes interaction with many molecules in cell, as well as, the single-stranded nature of RNA, which is normally involved in RNA recognition, due to the high base exposure and availability for interactions (Moore, 1999), AC could have a widespread application in RNA preparation (table 4).

Currently, the most widely and successfully used affinity chromatography strategy is the popular oligo(dT) for mRNA purification, based on the nucleotide base pairing specificity of A with T, since a poly(A) tail is virtually found in all eukaryotic mRNA molecules. These strategies allow mRNA purification directly from a biological sample or from previously isolated total RNA. mRNA enrichment is typically achieved using home-made columns or kits with oligo(dT)-cellulose (Aviv and Leder, 1972), but other oligo(dT)-derived media like biotinylated oligo(dT), streptavidin-coated latex or silica are also used (Kyte *et al.*, 2005, Ruby *et al.*, 1990). These methods are widely preferred among other isolation procedures because of the simplicity and reliable effectiveness in purifying mRNA molecules with high yields and purity suitable in cGMP production (Pascolo, 2008). However, because of particular interactions explored in these affinity techniques, the purification of non-polyadenylated mRNA, as exists in bacteria is not possible (Wendisch *et al.*, 2001).

RNA affinity tags have also emerged as useful tools for the isolation of RNAs from cell extracts in recent years (Ponchon and Dardel, 2011, Srisawat and Engelke, 2002). These tags are based on RNA aptamers, which have been selected for affinity against a specific ligand by using the systematic evolution of ligands by exponential enrichment (SELEX) method. A number of RNA aptamers have been used for the purification of RNAs from cellular extracts, including aptamers that bind to immobilized streptavidin or sephadex. These strategies use competitive elution with biotin or dextran, respectively, in order to recover the RNA of interest (Srisawat and Engelke, 2001, Srisawat *et al.*, 2001). Although SELEX permit the selection of practically all oligoribonucleotides sequences with the desired properties to bind to a specific molecule, the cost and availability of appropriate affinity resins where only those aptamers can bind is a huge limitation in the development of purification process based on affinity tags. In addition, several considerations for tagging the RNA to be purified, such as folding, steric blockage, and keeping the tag on the RNA prior to purification is required. Therefore, in some cases more than one purification step is required or denaturant conditions are employed in order to achieve higher RNA enrichments (Srisawat and Engelke, 2002, Srisawat *et al.*, 2001, Walker *et al.*, 2008).

In line with this, later investigations developed more robust affinity systems when pursuing the improvement of schemes to achieve rapid, large scale and native RNA purification of long constructs produced by *in vitro* transcription methods. These strategies are based on protein-RNA interactions where the affinity tags were attached to the 3′ end of the target RNAs during *in vitro* transcription, and specifically purified using an affinity column with a specific RNA-binding protein immobilized (Batey and Kieft, 2007, Kieft and Batey, 2004). Kieft and Batey described an affinity-based method where the RNA of interest was tagged by a signal recognition particle (SRP) RNA with the ability to bind to the immobilized RNA-binding domain of the SRP protein from *Tetrahymena thermophile*. Purification was achieved by cleavage with a *cis*-acting mutated ribozyme from the *hepatitis delta virus* (Kieft and Batey, 2004). In

a second strategy, the MS2 coat protein binding stem–loops was used as affinity tags at the 3' end of the RNA transcript, which binds to a Ni^{2+} affinity column via interaction with a hexahistidine-tagged MBP-MS2 coat protein fusion. This RNA tag is preceded by the cis-acting glmS ribozyme sequence which can be activated by glucosamine-6-phosphate to elute the desired RNA product (Batey and Kieft, 2007). These robust affinity purification methodologies allow the production of milligram quantities and the rapid purification of native RNAs with homogeneous 3'ends, which can be essential for crystallographic application. Although the major advantage of these systems is the broadly applicability to any RNA of interest, these methods are limited to the need of several design issues and binding the tags, which may lead to longer optimization processes. In addition, the use of ribozyme cleavage is not always suitable in purification procedures. The attached ribozyme may form alternative structures with the desired RNA and thus disrupt the correct folding of the ribozyme, with its self-cleaving power disabled (Sherlin *et al.*, 2001).

Recently, the use of amino acids as immobilized ligands for AC was exploited and implemented by our research group, as an effective methodology for the purification of RNA molecules (Martins *et al.*, 2010, Martins *et al.*, 2012b). These studies were based on the fact that many different specific interactions exist between proteins and nucleic acids in biological systems, involving in particular basic amino acids such as histidine or arginine (Sousa *et al.*, 2010). Moreover, several atomic and molecular recognition studies performed on RNA–protein interfaces have predicted preferential interactions occurring between particular amino acids and nucleotide bases (Jeong *et al.*, 2003, Treger and Westhof, 2001, Yarus *et al.*, 2009).

The use of less selective, but at the same time more robust, small molecular ligands (also known as pseudo bioaffinity ligands) was first introduced by Vijayalakshmi and coworkers for the purification of immunoglobulins and a wide variety of proteins, which has confirmed its exceptional potential (Vijayalakshmi, 1996). The concept of using these ligands was then applied by our research group for the first time to pDNA purification and their ability to isolate supercoiled (sc) pDNA proved the presence of specific interactions occurring between nucleic acids and the amino acid based matrices (Sousa *et al.*, 2010, Sousa *et al.*, 2008b).

Meanwhile, several studies are already available demonstrating the advantages of the amino acid-based affinity chromatography in the purification process of sc pDNA preparations intended for DNA vaccination or other gene therapy approaches (Sousa *et al.*, 2006, Sousa *et al.*, 2008a, Sousa *et al.*, 2009).

The potential of amino acids-based AC in obtaining RNA preparations under their native state and with high integrity and purity was proved in various biological RNA molecules from prokaryotic and eukaryotic cells rather than synthetic oligoribonucleotides. Although the use of biological RNAs has been diminished in many applications because of the improvements in

oligoribonucleotides synthesis, most of the native RNAs have post-transcriptional modifications (Dunin-Horkawicz *et al.*, 2006) that may not be reproduce under chemical or enzymatic synthesis and some of the modifications are quite important for their structure and function. Therefore, efficient methods for extraction and purification of RNA molecules from biological matrices are needed for the study of cellular mechanisms, as well as, in pharmacokinetic and pharmacodynamics analysis and biomarker evaluations of regulatory ncRNA (Wang *et al.*, 2008).

The potential applicability of histidine-based AC in the purification of RNA molecules was first demonstrated in the separation of 6S RNA, an ncRNA of the prokaryotic *E. coli* that has a relevant regulatory function in the transcription process. Because of the aromatic imidazole side chain of histidine amino acid the interactions between RNA and the matrix are suggested to be mainly hydrophobic either by ring-staking or hydrogen bonds, thus ammonium sulphate gradients were used in histidine strategies. Accordingly, a three stepwise decreasing gradient of ammonium sulphate was employed to histidine matrix that revealed a specific recognition for 6S RNA, allowing its accurate purification from a complex mixture of ncRNA molecules (Martins *et al.*, 2010). Subsequently, the simultaneous isolation of ncRNA and ribosomal RNA (rRNA) from *E.coli* cell lysates was also attained by an histidine-based method but in a single step gradient (Martins *et al.*, 2012b). In this study, both RNA classes were separated from host impurities (gDNA and proteins) showing a high purity degree and recovery yields comparable to conventional phenol-based methods, but with the advantage of avoiding the use of toxic chemicals during the isolation process. The versatility of histidine matrix in the purification of both single 6S RNA molecule and ncRNA and rRNA classes suggested that the underlying mechanism involves not only hydrophobic interaction, but also a bio-recognition of RNA bases by histidine (Martins *et al.*, 2010).

While the RNAs of these preparations need further functional characterization to prove their applicability in RNA-based studies, histidine methodology represents a major advance in the ability to accurately purify RNA molecules as the conventional procedures do not have the capacity to fractionate RNA at a single molecule and the simultaneous isolation of ncRNA and rRNA in a single procedure is not available yet. However, the requirement for high salt concentration can be a disadvantage, especially with regard to biotechnological application, because the use of salt is associated with higher costs and environmental impact.

The use of arginine as the immobilized ligand could improve the previous techniques. In fact, arginine-base interactions have been recognized as the most prevalent interactions in several protein-RNA complexes (Jeong *et al.*, 2003). As arginine is a positively charged amino acid, electrostatic interactions could be exploited for RNA purification using only mild elution conditions. Thus, a stepwise increasing gradient of sodium chloride up to 1M permitted the isolation of RNAt from impurities of eukaryotic cell extracts (Martins *et al.*, 2012a). Arginine support showed an exceptional ability to interact with all functional classes of RNA, despite

their structural diversity and different folding states as they are in their native state. These strong and selective interactions with the affinity matrix may result from the multiplicity of arginine side chain, which can promote multi contact with RNA backbone or RNA bases, according to RNA folding. Although the electrostatic interactions between RNA phosphate groups and the arginine ligands could be playing an important role on RNA retention, the bases contacts are also involved and modulate some favored interaction and specificity found in arginine–agarose chromatography. Hence, this process resulted in a high recovery yield of RNA and quality control analysis showed a high integrity in RNA preparations as well as good purity, demonstrated by the scarce detection of proteins and the reduction on genomic DNA contamination to residual concentrations. The efficiency and applicability of this technique was attested by gene expression analysis, where the RNAt samples were used as template in real-time PCR and the expression of two common housekeeping genes was successfully measured (Martins *et al.*, 2012a). Therefore, this arginine-based procedure can be an alternative method for the purification of total RNA pursuing gene expression analysis.

Taken together, the results on RNA isolation and purification from a complex biological mixture using amino acids-based AC techniques showed several improvements over the currently phenol/chloroform or SPE isolation methods, as it simplified the workflow integration and miniaturizes sample handling process, making them useful for the development of a RNase and organic solvents free methodology, particular important for several structural and functional studies and clinical investigation.

Table 4 - Affinity chromatography methods for RNA Purification. (Refs., References; cGMP, current good manufacturing practices; PAGE, polyacrylamide gel electrophoresis; SRP, signal recognition particle; ncRNA, non-coding RNA; rRNA, ribosomal RNA; SPE, solid phase extraction).

Affinity type	Principle	Advantages	Limitations	Refs.
Base pairing	Poly(A) tails present on mRNA molecules specifically interact with immobilized oligo(dT)	High mRNA yield cGMP quality Can be used without intermediate isolation of total RNA	Does not isolate no polyadenylated mRNA	Ruby *et al.* (1990)
Affinity tags	The RNA of interest is tagged with engineered oligoribonucleotides sequences (aptamers) that bind tightly to sephadex or streptavidin resins	Allow the isolation of RNA from cells lysates Robust methodologies using affinity tags based on inherent biologic interactions Stability of affinity tags High recovery yields due to the specific selection of the tags with high affinities for the resins	Sequential purifications for high recoveries Use of denaturing condition in recovery Several issue in the insertion of the affinity tag into RNA	Srisawat and Engelke (2002) Srisawat *et al.* (2001)
	Immobilized Tetrahymena thermophile M4 on a matrix or the commercial Ni^{2+}-affinity column specifically recognizes the affinity tags SRP RNA or MS2 coat protein binding stem-loops incorporated in the RNA of interest, respectively	Effective alternative to preparative PAGE in the purification of RNA prepared by *in vitro* transcription Robust methodologies using affinity tags based on inherent biologic interactions High recoveries	Time-demanding in design considerations Instability of the affinity tag	Batey and Kieft (2007), Kieft and Batey (2004)
Amino acids-RNA	Multiple interactions occur between immobilized histidine or arginine and RNA molecules	Isolates a single RNA type Simultaneous isolation of ncRNA and rRNA in a single procedure Effective alternative to phenol/chloroform and SPE purifications of biological total RNA Versatility on RNA purification Stable ligands High recovery yields	Requirement of high salt concentrations (for histidine)	Martins *et al.* (2010), Martins *et al.* (2012c)

In fact, these purification schemes can a be valuable tool in recent emerged strategies to produce stable structured RNA *in vivo* using recombinant approaches as an alternative to *in vitro* transcription or chemical synthesis. Since synthetic methods can often be costly and laborious and have their drawbacks with respect to sequence requirements, variations in yield, non-templated nucleotide additions and/or the maximum length of the oligonucleotide, tRNA-scaffolds have been described with some success as general method for producing a large variety of recombinant RNA using *E. coli* as a host, providing an inexhaustible source of RNA. (Ponchon and Dardel, 2011). At the end of the culture, cells are pelleted, lysed and RNAt is isolated by direct phenol extraction and the recombinant RNA is purified by AEC or alternatively, since the recombinant RNA molecules are equipped with a Sephadex aptamer, by AC using Sephadex G-200, avoiding phenol-based purification. Finally, the product RNA is separated from the tRNA-scaffold on preparative denaturing PAGE (Nelissen *et al.*, 2012). The use of amino-acid based methods to purify recombinant RNA would definitely improve the present of purification strategies, since the phenol/chloroform extraction as well as the preparative PAGE and all their disadvantages can be eliminated, maintaining the chemical stability and biological competence of RNA molecules.

In biopharmaceutical industry, the use of an appropriate recombinant expression system together with robust liquid chromatographic techniques is a well-established method for the production of therapeutic bioproducts, such as recombinant proteins or pDNA (Cart and Jungbauer, 2010, Ferreira *et al.*, 1999). Accordingly, the recombinant production of RNA in *E. coli* and the subsequent purification employing an amino acid-based AC method could be a promising biotechnological process for the production of RNA bioproducts in large scale. Figure 2 schematize a generic procedure for the preparation of recombinant RNA bioproduct based on amino acid-based AC technique. This approach could provide a cost-effective process, easily adopted by biopharmaceutical industries.

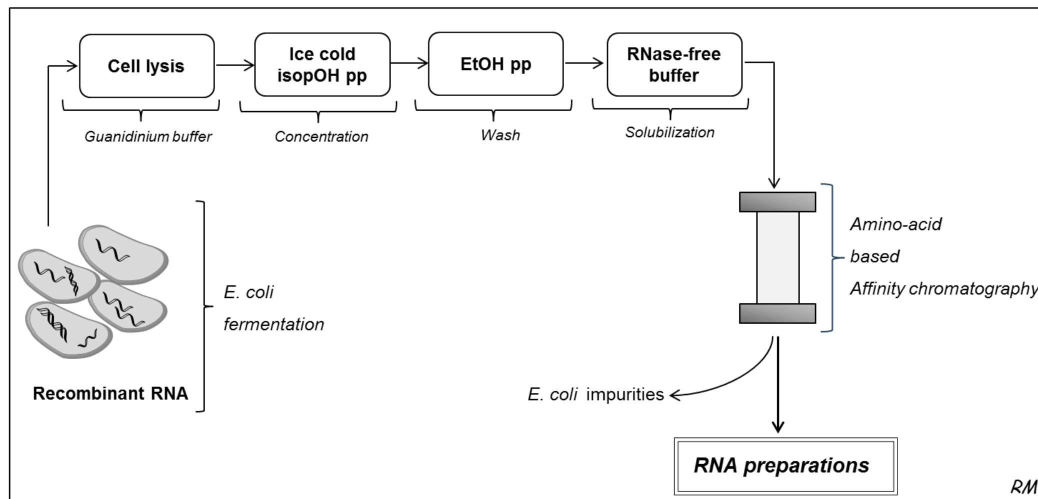

Figure 2 - **Projected biotechnological process for RNA preparation in large scale.** The downstream processing is based on previously reported amino acid-based affinity chromatography works (Martins *et al.*, 2012a, Martins *et al.*, 2012c). After amplification in *E. coli*, the recombinant RNA are lysed with a guanidinium-based buffer, followed by concentration, wash and solubilization steps in order to change to a simple buffer free of RNases, and a final step of amino acid-based affinity chromatography to isolate and purify RNA from impurities. (*E. coli*, *Escherichia coli*; RNases, Ribonuclease; IsopOH pp, isopropanol precipitation; EtOH pp, ethanol precipitation).

Whether the implementation of such method is feasible for the preparation of therapeutic RNA molecules, new high-throughput analytical tools that can readily identify and quantify RNA are required, since several flaws are still existing in the current RNA quantification techniques, as the lack of selectivity for RNA (Martins *et al.*, 2013). With the success of amino acids as affinity ligands in the purification of RNA, an analytical technique based on arginine AC for quantification and quality verification of RNAt was also developed and validated, according to international and European legislation for bioanalytical methods. The versatility of the methodology was demonstrated by its applicability in the quantification of RNA from different eukaryotic cells and in crude samples of chemically synthetized RNA, which can have a potential multipurpose applicability in molecular biology RNA-based analysis and RNA therapeutics (Martins *et al.*, 2013).

More efforts into RNA purification are being developed with amino acid based matrices, in particular with arginine-based matrix. The new goal is to exploit their applicability in purifying mRNA molecules not from cells, but from synthetic crudes of *in vitro* transcription, pursuing mRNA vaccination for cervical cancer. So far, mRNA molecules encoding human

papillomavirus (HPV) 16 E6 and E7 proteins were successfully purified from the impurities of pDNA template, enzymes, nucleotides, salts or buffer. In this work, arginine-based chromatography is also demonstrating its singular ability in improving purification processes, showing the advantages of eliminating additional steps and improving global economics of the production process (unpublished results).

5. Conclusion and future trends

To attain large quantities of RNA with the demanded quality for structural and functional studies as well as for application in medical therapies, more selective, reliable and efficient isolation and purification processes are required. The difficulty of successfully handling RNA in the laboratory, because of the structural susceptibility to RNases present in cell extracts and in the environment, and the impurities present in cell samples or synthetic preparations, result in numerous challenges for the selective purification of RNA molecules and in higher costs for laboratories and biopharmaceutical industries.

The current methodologies used in RNA preparations circumvent some of those challenges, however many difficulties still existing in order to obtain an innocuous, selective and economically feasible procedure to purify RNA. As several clinical trials are ongoing provided by many pharmaceutical and biotechnological companies, it seems that more accurate methods exists but are not ready accessible. Despite that, the motivation of the scientific community to disclose the knowledge on RNA purification processes is significantly increasing. Recent investigations in chromatographic bioseparation are providing reliable data in process development both for preparative-scale and for large-scale applications. In particular, AC stands out because of combining the selectivity of a naturally occurring biological interaction with the simplicity of a single small molecule. Moreover, the contribution of a novel affinity approach based on amino acid ligands provided new insights in the way that the diverse RNA molecules can be accurately purified. Moreover this methodology brings the prospect to be combined in a recombinant approach rising the opportunity for the development of a robust biotechnological process that might have a positive economic impact on the industrial production of RNA.

Although AC represents a particular attractive option for RNA purification, there are limitations regarding capacity of supports and the availability of suitable ligands that remain to be solved. However, with the increasing understanding on RNA synthetic synthesis and chemical modification, the design of highly selective and stable synthetic affinity ligands might be more easily achieved. In addition, recent chromatographic technologies as monolithic supports that combine the advantages of conventional chromatographic columns packed with porous particles in terms of separation power, capacity and sample distribution and those of membrane technology in terms of convective mass transport could be used with immobilized affinity ligands, which should largely benefit the industrial production of RNA biotherapeutics.

The combined efforts that are expected to occur in this field should help to accomplish enhanced purification strategies with positive effects on RNA basic and applied research.

Acknowledgments

This work was supported by FCT, the Portuguese Foundation for Science and Technology (PTDC/EBB-BIO/114320/2009, PEst-C/SAU/UI0709/2011 COMPETE and and EXPL/BBB-BIO/1056/2012, COMPETE: FCOMP-01-0124-FEDER-027560). Rita Martins also acknowledges a fellowship (SFRH/BD/ 64100/2009) from FCT.

The authors declared no conflict of interest.

References

Arraiano CM, Andrade JM, Domingues S, Guinote IB, Malecki M, Matos RG, et al. The critical role of RNA processing and degradation in the control of gene expression. FEMS Microbiol Rev. 2010; 34:883-923.

Aviv H, Leder P. Purification of Biologically Active Globin Messenger RNA by Chromatography on Oligothymidylic acid-Cellulose. Proc Natl Acad Sci U S A. 1972; 69:1408-12.

Ayyar BV, Arora S, Murphy C, O'Kennedy R. Affinity chromatography as a tool for antibody purification. Methods. 2012; 56:116-29.

Baker BF, Lot SS, Condon TP, Cheng-Flournoy S, Lesnik EA, Sasmor HM, et al. 2′-O-(2-Methoxy)ethyl-modified Anti-intercellular Adhesion Molecule 1 (ICAM-1) Oligonucleotides Selectively Increase the ICAM-1 mRNA Level and Inhibit Formation of the ICAM-1 Translation Initiation Complex in Human Umbilical Vein Endothelial Cells. J Biol Chem. 1997; 272:11994-2000.

Batey RT, Kieft JS. Improved native affinity purification of RNA. RNA. 2007; 13:1384-9.

Belostotsky R, Frishberg Y, Entelis N. Human mitochondrial tRNA quality control in health and disease: a channelling mechanism? RNA Biol. 2012; 9:33-9.

Blount KF, Breaker RR. Riboswitches as antibacterial drug targets. Nat Biotechnol. 2006; 24:1558-64.

Booy EP, Meng H, McKenna SA. Native RNA Purification by Gel Filtration Chromatography In: Conn GL, editor. Recombinant and In Vitro RNA Synthesis 2012. p. 69-81.

Brantl S. Antisense-RNA regulation and RNA interference. Biochim Biophys Acta. 2002; 1575:15-25.

Burnett JC, Rossi JJ. RNA-based therapeutics: current progress and future prospects. Chem Biol. 2012; 19:60-71.

Bustin SA, Nolan T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J Biomol Tech. 2004; 15:155-66.

Cart G, Jungbauer A. Downstream Processing of Biotechnology Products. Protein Chromatography Process Development and Scale-Up WILEY-VCH Verlag GmbH & Co.; 2010.

Chatterjee A, Mirer PL, Zaldivar Santamaria E, Klapperich C, Sharon A, Sauer-Budge AF. RNA isolation from mammalian cells using porous polymer monoliths: an approach for high-throughput automation. Anal Chem. 2010; 82:4344–56.

Chionh YH, Ho C-H, Pruksakorn D, Ramesh Babu I, Ng CS, Hia F, et al. A multidimensional platform for the purification of non-coding RNA species. Nucleic Acids Res. 2013.

Chomczynski P, Sacchi N. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. Nat Protoc. 2006; 1:581-5.

Cramer H, Finn KJ, Herzberg E. Purity Analysis and Impurities Determination by Reversed-Phase High-Performance Liquid Chromatography. In: Bonilla J.V., Srivatsa SG, editors. Handbook of Analysis of Oligonucleotides and Related Products. Boca Raton, Florida: CRC Press; 2011. p. 1-46.

Dias N, Stein CA. Antisense oligonucleotides: basic concepts and mechanisms. Mol Cancer Ther. 2002; 1:347-55.

Dickman MJ. Ion Pair Reverse-Phase Chromatography: A Versatile Platform for the Analysis of RNA. Chromatography Today 2011a. p. 22-6.

Dickman MJ. Ion Pair Reverse-Phase Chromatography: A Versatile Platform for the Analysis of RNA. Chromatography Today http://www.chromatographytoday.com 2011b. p. 22-6.

Directive 2001/83/EC. European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use: EC, Brussels, Belgium 2001 http://eceuropaeu/health/files/eudralex/vol-1/dir_2001_83_cons/dir2001_83_cons_20081230_enpdf.

Doench JG, Petersen CP, Sharp PA. siRNAs can function as miRNAs. Genes Dev. 2003; 17:438-42.

Doudna JA. Preparation of homogeneous ribozyme RNA for crystallization. Methods Mol Biol. 1997; 74:365-70.

Dunin-Horkawicz S, Czerwoniec A, Gajda MJ, Feder M, Grosjean H, Bujnicki JM. MODOMICS: a database of RNA modification pathways. Nucleic Acids Res. 2006; 34:D145-9.

Easton LE, Shibata Y, Lukavsky PJ. Rapid, nondenaturing RNA purification using weak anion-exchange fast performance liquid chromatography. RNA. 2010; 16:647-53.

El-Sagheer AH, Brown T. New strategy for the synthesis of chemically modified RNA constructs exemplified by hairpin and hammerhead ribozymes. Proc Natl Acad Sci U S A. 2010; 107:15329-34.

Farrell RE. RNA Methodologies. In: Meyers RA, editor. Epigenetic Regulation and Epigenomics: Advances in Molecular Biology and Medicine. 1st ed. Weinheim: Wiley-VCH; 2012. p. 1-37.

Ferreira GN, Cabral JM, Prazeres DM. Development of process flow sheets for the purification of supercoiled plasmids for gene therapy applications. Biotechnol Prog. 1999; 15:725-31.

Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. Mol Aspects Med. 2006; 27:126-39.

Ge Q, McManus MT, Nguyen T, Shen CH, Sharp PA, Eisen HN, et al. RNA interference of influenza virus production by directly targeting mRNA for degradation and indirectly inhibiting all viral RNA transcription. Proc Natl Acad Sci U S A. 2003; 100:2718-23.

Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet. 2009; 10:94-108.

Gilar M. Analysis and purification of synthetic oligonucleotides by reversed-phase high-performance liquid chromatography with photodiode array and mass spectrometry detection. Anal Biochem. 2001; 298:196-206.

Gjerde DT, Hoang L, Hornby D. RNA Purification and Analysis: Sample Preparation, Extraction, Chromatography: Wiley-VCH Verlag GmbH & Co; 2009.

Guidance for Industry: Consideration for plasmid DNA vaccines for infection disease indications. FDA, US 2007 http://wwwfdagov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/ucm091968pdf

Guideline on Human Cell-based Medicinal Products. EMA, L, Great Britain 2007 http://wwwemaeuropaeu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003898pdf.

Hagen FS, Young ET. Preparative polyacrylamide gel electrophoresis of ribonucleic acid. Identification of multiple molecular species of bacteriophage T7 lysozyme messenger ribonucleic acid. Biochemistry. 1974; 13:3394-400.

Jahn CE, Charkowski AO, Willis DK. Evaluation of isolation methods and RNA integrity for bacterial RNA quantitation. J Microbiol Methods. 2008; 75:318-24.

Jeong E, Kim H, Lee SW, Han K. Discovering the interaction propensities of amino acids and nucleotides from protein-RNA complexes. Mol Cells. 2003; 16:161-7.

Karikó K, Muramatsu H, Ludwig J, Weissman D. Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. Nucleic Acids Res. 2011; 39:e142.

Kieft JS, Batey RT. A general method for rapid and nondenaturing purification of RNAs. RNA. 2004; 10:988-95.

Kim I, McKenna SA, Viani Puglisi E, Puglisi JD. Rapid purification of RNAs using fast performance liquid chromatography (FPLC). RNA. 2007; 13:289-94.

Kiss T. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. EMBO J. 2001; 20:3617-22.

Knepp JH, Geahr MA, Forman MS, Valsamakis A. Comparison of automated and manual nucleic acid extraction methods for detection of enterovirus RNA. J Clin Microbiol. 2003; 41:3532-6.

Koubek J, Lin KF, Chen YR, Cheng RP, Huang JJT. Strong anion-exchange fast performance liquid chromatography as a versatile tool for preparation and purification of RNA produced by in vitro transcription. RNA. 2013.

Kreiter S, Diken M, Selmi A, Tureci O, Sahin U. Tumor vaccination using messenger RNA: prospects of a future therapy. Curr Opin Immunol. 2011; 23:399-406.

Kusenda B, Mraz M, Mayer J, Pospisilova S. MicroRNA biogenesis, functionality and cancer relevance. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2006; 150:205-15.

Kyte JA, Kvalheim G, Aamdal S, Saeboe-Larssen S, Gaudernack G. Preclinical full-scale evaluation of dendritic cells transfected with autologous tumor-mRNA for melanoma vaccination. Cancer Gene Ther. 2005; 12:579-91.

Kyte JA, Mu L, Aamdal S, Kvalheim G, Dueland S, Hauser M, et al. Phase I/II trial of melanoma therapy with dendritic cells transfected with autologous tumor-mRNA. Cancer Gene Ther. 2006; 13:905-18.

Labrou NE. Design and selection of ligands for affinity chromatography. J Chromatogr B. 2003; 790:67-78.

Laurell H, Iacovoni JS, Abot A, Svec D, Maoret J-J, Arnal J-F, et al. Correction of RT-qPCR data for genomic DNA-derived signals with ValidPrime. Nucleic Acids Res. 2012.

Lehman N. RNA in evolution. Wiley Interdiscip Rev RNA. 2010; 1:202-13.

Lin SL, Miller JD, Ying SY. Intronic microRNA (miRNA). J Biomed Biotechnol. 2006; 2006:26818.

Loens K, Bergs K, Ursi D, Goossens H, Ieven M. Evaluation of NucliSens easyMAG for automated nucleic acid extraction from various clinical specimens. J Clin Microbiol. 2007; 45:421-5.

Lorenzi J, Trombone A, Rocha C, Almeida L, Lousada R, Malardo T, et al. Intranasal vaccination with messenger RNA as a new approach in gene therapy: Use against tuberculosis. BMC Biotechnol. 2010; 10:77.

Lowe CR, Lowe AR, Gupta G. New developments in affinity chromatography with potential application in the production of biopharmaceuticals. J Biochem Biophys Methods. 2001; 49:561-74.

Lukavsky PJ, Puglisi JD. Large-scale preparation and purification of polyacrylamide-free RNA oligonucleotides. RNA. 2004; 10:889-93.

Magdeldin S, MoserTech A. Affinity Chromatography: Principles and Applications. In: Magdeldin S, editor. Affinity Chromatography: InTech; 2012.

Maina A, Blackman BA, Parronchi CJ, Morozko E, Bender ME, Blake AD, et al. Solid-phase synthesis, characterization and RNAi activity of branch and hyperbranch siRNAs. Bioorg Med Chem Lett. 2013; 23:5270-4.

Maroney PA, Chamnongpol S, Souret F, Nilsen TW. A rapid, quantitative assay for direct detection of microRNAs and other small RNAs using splinted ligation. RNA. 2007; 13:930-6.

Martins R, Maia CJ, Queiroz JA, Sousa F. A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography. J Sep Sci. 2012a; 35:3217-26.

Martins R, Queiroz JA, Sousa F. A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography. J Mol Recognit. 2010; 23:519-24.

Martins R, Queiroz JA, Sousa F. Histidine affinity chromatography-based methodology for the simultaneous isolation of Escherichia coli small and ribosomal RNA. Biomed Chromatogr. 2012b; 26:781-8.

Martins R, Queiroz JA, Sousa F. New approach in RNA quantification using arginine-affinity chromatography: potential application in eukaryotic and chemically synthesized RNA. Anal Bioanal Chem. 2013: 405:8849-58.

Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol. 2007; 8:209-20.

Mayer G. The chemical biology of aptamers. Angew Chem Int Ed Engl. 2009; 48:2672-89.

McGinnis AC, Chen B, Bartlett MG. Chromatographic methods for the determination of therapeutic oligonucleotides. J Chromatogr B Analyt Technol Biomed Life Sci. 2012; 883-884:76-94.

McGookin R. Purification of Synthetic Oligonucleotides by Preparative Gel Electrophoresis. 1988. p. 215-20.

McKenna SA, Kim I, Puglisi EV, Lindhout DA, Aitken CE, Marshall RA, et al. Purification and characterization of transcribed RNAs using gel filtration chromatography. Nat Protoc. 2007; 2:3270-7.

Melnikova I. RNA-based therapies. Nat Rev Drug Discov. 2007; 6:863-4.

Moore PB. Structural motifs in RNA. Annu Rev Biochem. 1999; 68:287-300.

Murray RK, Granner DK, Mayes PA, Rodwell VW. Harper´s Biochemistry. 25 ed. USA: McGraw-Hill; 2000.

Nelissen FH, Leunissen EH, van de Laar L, Tessari M, Heus HA, Wijmenga SS. Fast production of homogeneous recombinant RNA--towards large-scale production of RNA. Nucleic Acids Res. 2012; 40:e102.

Ni X, Castanares M, Mukherjee A, Lupold SE. Nucleic acid aptamers: clinical applications and promising new horizons. Curr Med Chem. 2011; 18:4206-14.

Nilsen TW. Selective Precipitation of Large RNAs. Cold Spring Harb Protoc. 2012; 2012:pdb.prot072322.

Noll B, Seiffert S, Hertel F, Debelak H, Hadwiger P, Vornlocher HP, et al. Purification of small interfering RNA using nondenaturing anion-exchange chromatography. Nucleic Acid Ther. 2011; 21:383-93.

Pascolo S. Vaccination with Messenger RNA (mRNA). In: Bauer S, Hartmann G, editors. Toll-Like Receptors (TLRs) and Innate Immunity: Springer Berlin Heidelberg; 2008. p. 221-35.

Phylactou LA, Kilpatrick MW, Wood MJ. Ribozymes as therapeutic tools for genetic disease. Hum Mol Genet. 1998; 7:1649-53.

Ponchon L, Dardel F. Large scale expression and purification of recombinant RNA in Escherichia coli. Methods. 2011; 54:267-73.

Price SR, Ito N, Oubridge C, Avis JM, Nagai K. Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. J Mol Biol. 1995; 249:398-408.

Raam BJv, Salvesen GS. Transferring Death: A Role for tRNA in Apoptosis Regulation. Mol Cell. 2010; 37:591-2.

Romanovskaya A, Sarin LP, Bamford DH, Poranen MM. High-throughput purification of double-stranded RNA molecules using convective interaction media monolithic anion exchange columns. J Chromatogr A. 2013; 1278:54-60.

Roque AC, Lowe CR. Affinity chromatography: history, perspectives, limitations and prospects. Methods Mol Biol. 2008; 421:1-21.

Ruby SW, Goelz SE, Hostomsky Z, Abelson JN. Affinity chromatography with biotinylated RNAs. Methods Enzymol. 1990; 181:97-121.

Scott LG, Hennig M. RNA structure determination by NMR. Methods Mol Biol. 2008; 452:29-61.

Shanagar J. Purification of a synthetic oligonucleotide by anion exchange chromatography: Method optimisation and scale-up. J Biochem Biophys Methods. 2005; 64:216-25.

Sharp PA. The Centrality of RNA. Cell. 2009; 136:577-80.

Sherlin LD, Bullock TL, Nissan TA, Perona JJ, Lariviere FJ, Uhlenbeck OC, et al. Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. RNA. 2001; 7:1671-8.

Soll D, Nishimura S, Moore P. RNA. First ed. Oxford: Pergamon; 2001.

Sousa F, Cruz C, Queiroz JA. Amino acids-nucleotides biomolecular recognition: from biological occurrence to affinity chromatography. J Mol Recognit. 2010; 23:505-18.

Sousa F, Freitas S, Azzoni AR, Prazeres DM, Queiroz JA. Selective purification of supercoiled plasmid DNA from clarified cell lysates with a single histidine-agarose chromatography step. Biotechnol Appl Biochem. 2006; 45:131-40.

Sousa F, Matos T, Prazeres DM, Queiroz JA. Specific recognition of supercoiled plasmid DNA in arginine affinity chromatography. Anal Biochem. 2008a; 374:432-4.

Sousa F, Prazeres DM, Queiroz JA. Affinity chromatography approaches to overcome the challenges of purifying plasmid DNA. Trends Biotechnol. 2008b; 26:518-25.

Sousa F, Prazeres DMF, Queiroz JA. Improvement of transfection efficiency by using supercoiled plasmid DNA purified with arginine affinity chromatography. J Gene Med. 2009; 11:79-88.

Srisawat C, Engelke DR. Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. RNA. 2001; 7:632-41.

Srisawat C, Engelke DR. RNA affinity tags for purification of RNAs and ribonucleoprotein complexes. Methods. 2002; 26:156-61.

Srisawat C, Goldstein IJ, Engelke DR. Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures. Nucleic Acids Res. 2001; 29:e4.

Srivatsa GS. Regulatory Considerations for the Development of Oligonucleotide Therapeutics. In: Bonilla JV, Srivatsa SG, editors. Handbook of Analysis of Oligonucleotides and Related Products. Boca Raton, Florida: CRC Press 2011. p. 465–86.

Srivatsa S. Regulatory Challenges. European Biopharmaceutical Review. 2012; Sping 2012.

Stillian JR, Pohl CA. New latex-bonded pellicular anion exchangers with multi-phase selectivity for high-performance chromatographic separations. J Chromatogr A. 1990; 499:249-66.

Sullenger BA, Gilboa E. Emerging clinical applications of RNA. Nature. 2002; 418:252-8.

Tenson T, Mankin A. Antibiotics and the ribosome. Mol Microbiol. 2006; 59:1664-77.

Thayer J, Murugaiah V., WU Y. Purity Analysis and Impurities Determination by AEX-HPLC. In: Bonilla J.V., Srivatsa SG, editors. Handbook of Analysis of Oligonucleotides and Related Products. Boca Raton, Florida: CRC Press; 2011a. p. 47–104.

Thayer JR, Flook KJ, Woodruff A, Rao S, Pohl CA. New monolith technology for automated anion-exchange purification of nucleic acids. J Chromatogr B. 2010; 878:933-41.

Thayer JR, Rao S, Pohl C, Wu Y, Angelino M. Pellicular anion-exchange chromatography applied to RNAi assays for monitoring strand stoichiometry and RNA stability. Chromatography Today http://www.chromatographytoday.com 2011b. p. 28-32.

Treger M, Westhof E. Statistical analysis of atomic contacts at RNA-protein interfaces. J Mol Recognit. 2001; 14:199-214.

Tucker BJ, Breaker RR. Riboswitches as versatile gene control elements. Curr Opin Struct Biol. 2005; 15:342-8.

Uhlen M. Affinity as a tool in life science. Biotechniques. 2008; 44:649-54.

Vijayalakshmi MA. Histidine Ligand Affinity Chromatography. Mol Biotechnol. 1996; 6:347-57.

Vomelova I, Vanickova Z, Sedo A. Methods of RNA purification. All ways (should) lead to Rome. Folia Biol (Praha). 2009; 55:243-51.

Walker SC, Scott FH, Srisawat C, Engelke DR. RNA affinity tags for the rapid purification and investigation of RNAs and RNA-protein complexes. Methods Mol Biol. 2008; 488:23-40.

Wang WX, Wilfred BR, Baldwin DA, Isett RB, Ren N, Stromberg A, et al. Focus on RNA isolation: obtaining RNA for microRNA (miRNA) expression profiling analyses of neural tissue. Biochim Biophys Acta. 2008; 1779:749-57.

Warren W, Vella G. Principles and methods for the analysis and purification of synthetic deoxyribonucleotides by high-performance liquid chromatography. Molecular Biotechnology. 1995; 4:179-99.

Weide B, Garbe C, Rammensee HG, Pascolo S. Plasmid DNA- and messenger RNA-based anti-cancer vaccination. Immunol Lett. 2008; 115:33-42.

Wen J, Legendre LA, Bienvenue JM, Landers JP. Purification of Nucleic Acids in Microfluidic Devices. Anal Chem. 2008; 80:6472-9.

Wendisch VF, Zimmer DP, Khodursky A, Peter B, Cozzarelli N, Kustu S. Isolation of *Escherichia coli* mRNA and comparison of expression using mRNA and total RNA on DNA microarrays. Anal Biochem. 2001; 290:205-13.

Wittmann A, Suess B. Engineered riboswitches: Expanding researchers' toolbox with synthetic RNA regulators. FEBS Lett. 2012; 586:2076-83.

Yarus M, Widmann JJ, Knight R. RNA-amino acid binding: a stereochemical era for the genetic code. J Mol Evol. 2009; 69:406-29.

Yu RZ, Zhang H, Geary RS, Graham M, Masarjian L, Lemonidis K, et al. Pharmacokinetics and Pharmacodynamics of an Antisense Phosphorothioate Oligonucleotide Targeting Fas mRNA in Mice. J Pharmacol Exp Ther. 2001; 296:388-95.

Zeng Y, Wagner EJ, Cullen BR. Both Natural and Designed Micro RNAs Can Inhibit the Expression of Cognate mRNAs When Expressed in Human Cells. Mol Cell. 2002; 9:1327-33.

Chapter 2

Aims of the thesis

The main goal of this thesis is to explore the possibility of using affinity chromatography with amino acids as specific ligands to develop new methodologies to efficiently isolate, purify and analyze RNA molecules. Because of the particularities of RNA and the challenges implicit to its purification, the interactions between amino acids and RNA will be exploited so that RNA molecules can be recovered under their native state and with high integrity and purity, in order to follow the growing demands of the therapeutic application of RNA. Since various limitations still exist in the current techniques to isolate and purify RNA, namely the use of chemical solvents and denaturing conditions, the use of amino acid supports in affinity chromatography is intended to aid in the development of new approaches to recover a product to be biologically applied and that lead to an efficient, tolerable and scalable process, improving workload, process time and economics. To achieve this principal purpose, the work will be developed concerning the following specific aims:

1. Optimization of the lysis and separation processes to recover and isolate RNA molecules or total nucleic acids from prokaryotic or eukaryotic cells, for application in chromatographic studies;

2. Study of the binding/elution conditions of RNA or total nucleic acids extracts in affinity supports with immobilized histidine or arginine by testing the influence of elution conditions such as the temperature, flow rate and buffer composition (pH, ionic strength and type of salt), in order to evaluate the ability of the supports to specifically recognize the various RNA molecules.

3. Monitoring the influence of the chromatographic conditions in the structure and stability of RNA and control its quality throughout the purification process;

4. Development of purification strategies to selectively purify RNA molecules from complex and impure cell extracts in order to eliminate the use of organic solvents and denaturing conditions from RNA isolations;

5. Evaluation of the biological activity of the purified RNAs by integrating the chromatographic process with the more reliable and promising support in the global process of RNA preparation to be applied in RNA research or RNA therapeutics.

Chapter 3

Paper II

A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography

R. Martins, J. A. Queiroz, F. Sousa

Journal of Molecular Recognition. 2010. 23(6): 519-524

Received: 13 November 2009, Revised: 19 July 2010, Accepted: 20 July 2010, Published online in Wiley Online Library: 2010

(wileyonlinelibrary.com) DOI:10.1002/jmr.1078

A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography†

R. Martins[a], J. A. Queiroz[a] and F. Sousa[a]*

6S RNA is an abundant non-coding RNA in *Escherichia coli* (*E. coli*), but its function has not been discovered until recently. The first advance on 6S RNA function was the demonstration of its ability to bind the σ^{70}-holoenzyme form of RNA polymerase, inhibiting its activity and consequently the transcription process. The growing interest in the investigation of non-coding small RNAs (sRNA) calls for the development of new methods for isolation and purification of RNA. This work presents an optimized RNA extraction procedure and describes a new affinity chromatography method using a histidine support to specifically purify 6S RNA from other *E. coli* sRNA species. The RNA extraction procedure was optimized, and a high yield was obtained in the separation of sRNA and ribosomal RNA (rRNA) from total RNA (RNAt). This improved method takes advantage of its simplicity and significant cost reduction, since some complex operations have been eliminated. A purification strategy was also developed to separate 6S RNA from an sRNA mixture. Pure RNA can be advantageously obtained using the histidine-affinity chromatography method, aiming at its application to structural or functional studies. Copyright © 2010 John Wiley & Sons, Ltd.

Keywords: affinity chromatography; histidine; RNA; RNA extraction; small RNA; 6S RNA

INTRODUCTION

Small RNAs (sRNA) are a family of regulatory non-coding RNAs, which are targeted in several research studies on all organisms. Their activities have been associated to RNA processing, mRNA stability, translation, protein stability and secretion in bacteria, acting mostly by base pairing and, to a lesser extent, by protein binding (Wassarman *et al.*, 1999). *Escherichia coli* (*E. coli*) 6S RNA was one of the first sRNA to be discovered in the late 1960s, its sequence and secondary structure having been promptly proposed (Willkomm and Hartmann, 2005). Although 6S RNA was soon recognized, its function remained unknown for several years. This RNA species has always been found at high abundance in total cellular RNA (RNAt). However, studies of 6S RNA depletion and over-expression have not demonstrated any evident change in cell metabolism behaviour (Willkomm and Hartmann, 2005). In 2000, the first breakthrough on 6S RNA function was the demonstration of its ability to bind the σ^{70}-holoenzyme form of RNA polymerase, inhibiting its activity in the stationary phase of cell growth. The high structural similarity observed between 6S RNA and a DNA open promoter in RNA transcription process, as well as the identification of 6S RNA in a specific complex linked to RNA polymerase, has suggested that this RNA type functions by mimicking the transcription promoter. Hence, 6S RNA seems to directly compete with DNA promoters for the σ^{70}-RNA polymerase active site to inhibit the transcription process (Wassarman and Storz, 2000; Wassarman, 2007).

Over the last years, new insights triggered a better understanding of the molecular mechanism of 6S RNA (Gildehaus *et al.*, 2007; Karen, 2007), while the techniques employed in the isolation of this RNA species were still plasmid design and enzymatic purification. These techniques involve time-consuming and expensive procedures. Therefore, the development of new tools for 6S RNA purification would be of great importance to simplify the current genetic-based approaches.

The use of amino acid-based affinity chromatography has been described as a promising approach for nucleic acids purification. This is due to the biological selectivity, which occurs between the amino acid ligand and the nucleic acid molecule under study. In general, the selectivity found in affinity chromatography can be explained by some biological recognition or individual chemical structure, which favours the interaction (Sousa *et al.*, 2008). In the present study, the isolated *E. coli* sRNA population was tested in an affinity chromatographic support with immobilized histidine, combining the reliable and economical chromatographic operation with the high selectivity of histidine, which has been shown to have great applicability in the purification of nucleic acids. Histidine–agarose chromatographic support has demonstrated to efficiently separate supercoiled (SC) and open circular (OC) plasmid DNA (pDNA) isoforms, via a bio-recognition phenomenon with the nucleic acid bases, involving also hydrophobic interactions between the support and the pDNA molecules (Sousa *et al.*, 2005). Furthermore, in these previous studies, the RNA showed to be strongly retained on the column, because of the higher base exposure, and a decrease in salt concentration was deemed necessary for its elution (Sousa *et al.*, 2006). Building on the interesting results obtained for RNA separation, this report explores the possibility of using histidine–agarose chromatography to purify 6S RNA from other sRNA.

* *Correspondence to: F. Sousa, CICS – Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, 6201-001 Covilhã, Portugal. E-mail: fani.sousa@fcsaude.ubi.pt*

a *R. Martins, J. A. Queiroz, F. Sousa*
CICS – Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, 6201-001 Covilhã, Portugal

† *This article is published in Journal of Molecular Recognition as a special issue on Affinity 2009, edited by Gideon Fleminger, Tel-Aviv University, Tel-Aviv, Israel and George Ehrlich, Hoffmann-La Roche, Nutley, NJ, USA.*

Copyright © 2010 John Wiley & Sons, Ltd.

MATERIALS AND METHODS

Materials

L-histidine–agarose gel, oligo(dT) cellulose and all the organic compounds used in the RNAt extractions were obtained from Sigma (St. Louis, MO, USA). Glycogen was from Roche (Mannheim, Germany) and the RNA/DNA midi kit was purchased from Qiagen (Hilden, Germany). The PD-10 desalting columns were from GE Healthcare (Uppsala, Sweden) and the RNA molecular weight marker was obtained from Invitrogen (Carlsbad, CA, USA). Specific primers for *E. coli* 6S RNA cDNA were purchased from Stab Vida (Lisbon, Portugal). All salts used were of analytical grade.

Bacterial growth conditions

The RNA used in this study was obtained from a cell culture of *E. coli* DH5α. Growth was carried out in shake flasks at 37°C and 250 rpm with 250 mL of Terrific Broth medium (12 g/L tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 0.017 M KH_2PO_4, 0.072 M K_2HPO_4). It was suspended in the late log phase ($OD_{600} \approx 9$). Cells were recovered by centrifugation and were stored at −20°C.

Lysis and RNAt isolation

Cells were lysed and RNAt was extracted using the acid guanidinium thiocyanate–phenol–chloroform method described by Chomczynski and Sacchi (2006) with some modifications. The bacterial pellets were resuspended in 5 mL of Solution D (4 M guanidinium thiocyanate; 25 mM sodium citrate, pH 4.0; 0.5% (w/m) *N*-laurosylsarcosine (sarcosyl) and 0.1 M β-mercaptoethanol) to perform lysis. After incubating on ice for 10 min, cellular debris, genomic DNA and proteins were precipitated by gently adding and mixing 10 mL of water-saturated phenol and 1 mL of 2 M sodium acetate (pH 4.0). The RNAt isolation was achieved by adding 2 mL of chloroform/isoamyl alcohol (49:1), and by mixing vigorously until two immiscible phases were obtained. The upper aqueous phase, which contained mostly RNA, was recovered and concentrated by the addition of 11 μL of 20 μL/mL glycogen and 10 mL of ice-cold isopropanol. After incubating for 10 min at −20°C, RNAt was recovered by centrifugation at 15 000 *g* for 20 min at 4°C. The RNAt pellet was resuspended in 3 mL of Solution D. It was concentrated again with glycogen and 3 mL of ice-cold isopropanol. After centrifuging for 10 min at 15 000 *g* (4°C), the RNAt pellet was washed with 7.5 mL of 75% ethanol and incubated at room temperature for 10 min, followed by a 5 min centrifugation at 15 000 *g* (4°C). The air-dried RNAt pellet was solubilized in 400 μL of 0.05% DEPC-treated water, and its optical density was determined in order to assess RNAt quantity and purity.

Isolation of sRNA population from RNAt extract

The isolation of the sRNA population was first performed using the method previously described by Stellrecht and Gandhi (2002). However, the complexity of this protocol and its frequent use in eucaryotic cells led us to develop a new methodology for sRNA extraction from procaryotic cells.

In the optimized protocol, RNAt extract was precipitated with 2 M $(NH_4)_2SO_4$ and incubated at 4°C for several hours. The sRNA recovery was accomplished by centrifugation at 15 000 *g* for 20 min at 4°C, as the sRNA population was in the supernatant, while the ribosomal RNA (rRNA) was in the pellet. Next, the supernatant was applied to a PD-10 desalting column (GE Healthcare, Uppsala, Sweden), following the manufacturer's instructions. The sRNA concentration was achieved with 2 volumes of ice-cold isopropanol and 11 μL of 20 μL/mL glycogen, followed by centrifugation at 15 000 *g* for 20 min at 4°C. Finally, the sRNA pellet was solubilized with 200 μL of 0.05% DEPC-treated water. The quantification and purity of sRNA was assessed as previously described for RNAt.

Affinity chromatography

Chromatography was performed in an ÄKTA purifier system with UNICORN software (GE Healthcare, Sweden). A 10 mm diameter × 20 mm long (~2 mL) column was packed with the commercial L-histidine–agarose gel. The manufacturer characterizes this support as a cross-linked 4% beaded agarose matrix with a 12-atom spacer and an extent of labelling between 1 and 2 μmol/mL. The column was equilibrated with 2.2 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 7.0) buffer at a flow rate of 1 mL/min. The isolated sRNA samples were injected onto the column using a 100 μL loop at the same flow rate. The absorbance of the eluent was continuously monitored at 260 nm. After elution of unbound species with 2.2 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 7.0) buffer, the ionic strength of the buffer was decreased stepwise to 1.6 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 7.0) buffer. Finally, tightly bound RNA species were removed by changing to ammonium sulphate-free 10 mM Tris–HCl (pH 7.0) buffer. Fractions were pooled according to the chromatograms obtained. Following concentration and desalting with Vivaspin concentrators (Vivascience), the pools were kept for further analysis as described below.

Polyacrylamide electrophoresis

Pooled fractions were analysed by vertical electrophoresis using Amersham Biosciences system (GE Healthcare, Sweden) with 7.5% polyacrylamide gel and were then stained with ethidium bromide (0.5 μg/mL). Electrophoresis was carried out at 120 V for 90 min with TBE buffer (0.84 M Tris base, 0.89 M boric acid and 0.01 M EDTA, pH 8.3). The pooled fractions were previously denatured with 97.5% formamide and the denatured conditions were kept in the gel due to the presence of 8 M urea. sRNA in the gel was visualized using a Vilber Lourmt system (ILC Lda).

Reverse-transcription PCR analysis

6S RNA identification was assessed using reverse-transcriptase polymerase chain reaction (PCR) in a MyCycler^{TM}Thermal Cycle (Biorad). Samples collected after the chromatographic purification process with histidine–agarose column were pre-treated with DNase I (Sigma Chemical Co., St. Louis, MO) according to the manufacturer's instructions. cDNA was synthesized from 150 or 500 ng of 6S RNA which was denatured for 5 min at 65°C with 500 μM of deoxynucleotide triphosphates (Amersham, Uppsala, Sweden) and 250 ng of random primers (Invitrogen, Karlsruhe LMA, Germany). Reverse transcription was carried out at 37°C for 60 min in a 20 μL reaction containing reverse transcriptase buffer (50 mM Tris–HCl, 75 mM KCl and 3 mM $MgCl_2$), 0.1 M DTT, 60 U of RNaseOUT (Invitrogen) and 200 U of M-MLV RT (Invitrogen). The reaction was stopped at 75°C for 15 min. PCR reactions were carried out using 1 μL of synthesized cDNAs in a 25 μL reaction containing 1× Taq DNA polymerase buffer (20 mM Tris–HCl and 50 mM KCl), 500 μM deoxynucleotide triphosphates (Amersham),

Copyright © 2010 John Wiley & Sons, Ltd.

3 mM of $MgCl_2$ (Promega, Madison), 300 nM of each primer and 0.125 U of Taq DNA polymerase (Promega). Specific primers for *E. coli* 6S RNA cDNA (sense: 5′-GCT CCG CGG TTG GTG AGC AT-3′; antisense: 5′-GAT GCC GCC GCA GGC TGT AA-3′) whose design was achieved on RNA database were used to amplify a fragment of 95 bp. After an initial denaturation at 95°C for 5 min, the cycling conditions were used as follows: 25 cycles consisting of denaturation at 95°C for 30 s, annealing at 60°C for 30 s and extension at 72°C for 10 s. PCR products were analysed using 1.5% agarose-gel electrophoresis.

RESULTS AND DISCUSSION

The available protocols to obtain sRNA molecules are complex, time-consuming and require the use of highly toxic compounds. Therefore, they may not be adequate to recover a product for biological application, which justifies the need for the development of an efficient, tolerable and scalable process to purify non-coding RNA. This study was developed to evaluate the applicability of amino acid-based affinity chromatography, namely the potential of histidine ligand, in the purification of the different sRNA molecules.

Optimization of the RNA extraction methods

The methods commonly used to isolate the different types of RNA require a variety of organic compounds and saline solutions, as well as the application of commercial kits. In fact, this may be considered a drawback to their broader application. However, the developed protocol allowed the efficient isolation of RNA while significantly reducing the number of operation steps.

Figure 1 shows the global extraction procedure used in our experiments to obtain sRNA molecules while Figure 2 highlights the differences between the described multi-step method (Stellrecht and Gandhi, 2002) and the optimized protocol that we have developed to isolate the sRNA population from RNAt.

The herein proposed protocol was based on the stringent properties of RNA, consisting of the manipulation of several variables, such as the ionic strength, which influence the nucleic acid structure (Farrell, 2005). The effect of high salt concentrations on stringency allows single-stranded nucleic acid molecules to form stable hydrogen bonds between their complementary bases. The monovalent cations present in the salts minimize the tendency for natural electrostatic repulsion between two negatively charged phosphodiester backbones and, as a result, the nucleic acid structure is stabilized and compacted. In particular, salts with NH_4^+ cations stabilize the tertiary structure of rRNA due to the formation of specific bonds involving hydrogen bonds with a tetrahedral rearrangement of the rRNA carbonyl bases (Wang *et al.*, 1993). Thus, the separation of sRNA molecules was achieved via the precipitation of rRNA in the RNAt extract with 2 M $(NH_4)_2SO_4$ at 4°C for several hours. This procedure revealed to be practical, economical and effective in the isolation of the sRNA population for chromatographic application.

Chromatographic separation of sRNA 6S

The ability of the histidine–agarose support to separate the SC and OC pDNA isoforms present in a 'native' (SC + OC) pDNA sample prepared with a commercial purification kit or from a *E. coli* lysate has been recently described (Sousa *et al.*, 2005; Sousa *et al.*, 2006). The previous works demonstrated that RNA interacts

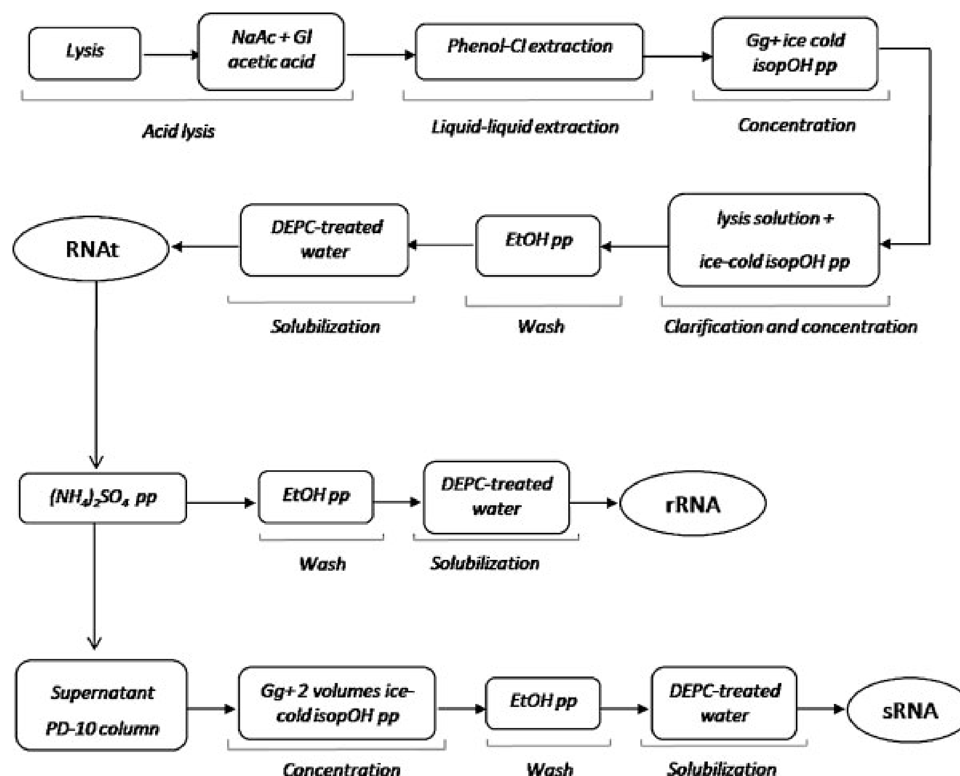

Figure 1. RNAt extraction and isolation of rRNA and sRNA (NaAc, sodium acetate; G, glacial; CI, chloroform/isoamyl alcohol; Gg, glycogen; IsopOH, isopropanol; etOH, ethanol; DEPC, diethylpyrocarbonate; $(NH_4)_2SO_4$, ammonium sulphate; pp, precipitation).

 Copyright © 2010 John Wiley & Sons, Ltd.

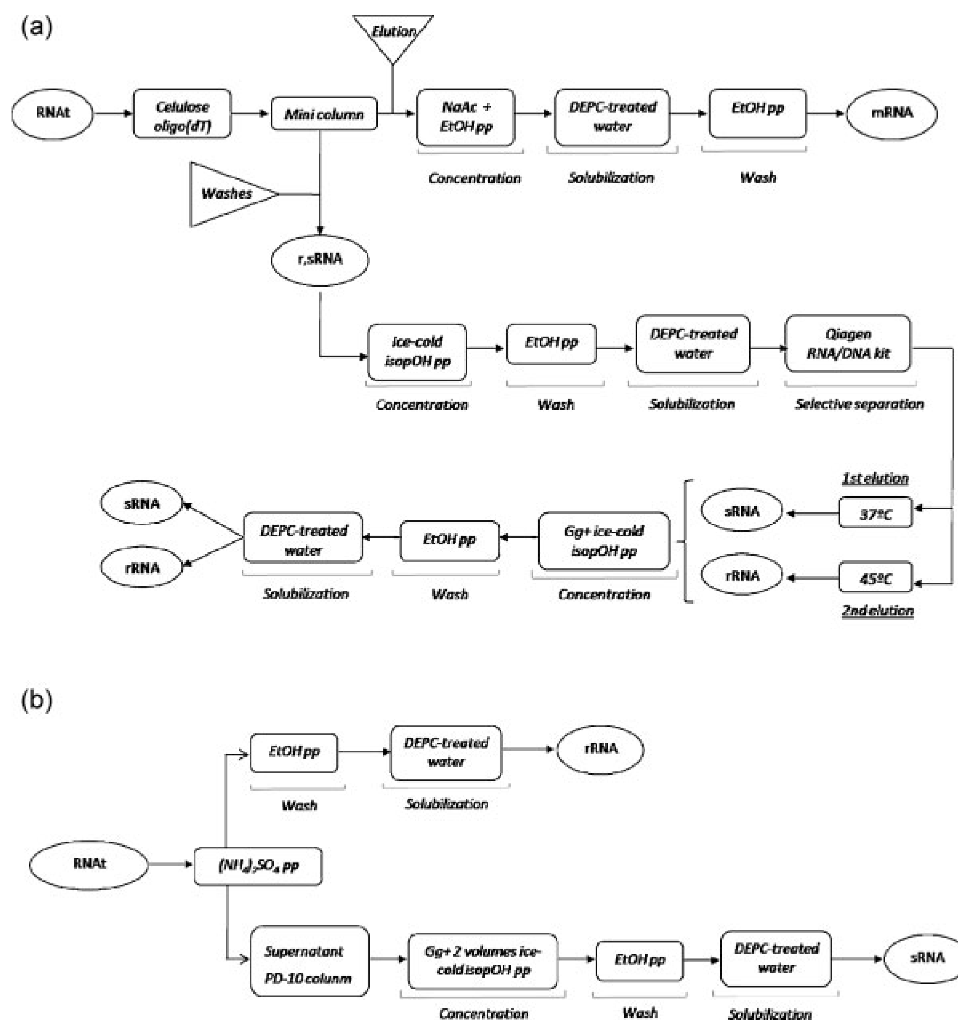

Figure 2. (a) Separation of mRNA, rRNA and sRNA using the method previously described by Stellrecht and Gandhi (b) rRNA and sRNA separation using the optimized protocol (NaAc, sodium acetate; G, glacial; CI, chloroform/isoamyl alcohol; Gg, glycogen; IsopOH, isopropanol; etOH, ethanol; DEPC, diethylpyrocarbonate; $(NH_4)_2SO_4$, ammonium sulphate; pp, precipitation).

with the histidine–agarose matrix in the presence of higher salt concentration and elutes with the decrease of $(NH_4)_2SO_4$ concentration in the elution buffer to 1.5 M. Following these results, the general aim of the present study was to explore the retention pattern of the different sRNA in the histidine–agarose support in order to investigate the interactions that might arise from RNA retention.

The 6S RNA purification process started with the extraction procedure for RNAt followed by the isolation of the sRNA species, which were required to perform the chromatographic studies in the histidine–agarose matrix.

The chromatographic assays performed initially were intended to test the ionic strength effects on the sRNA retention (data not shown). In the first set of experiments, a two-step elution with 2.2 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 7.0) buffer and ammonium sulphate-free Tris buffer was carried out. The higher salt concentration promoted total sRNA retention. Its elution was achieved with ammonium sulphate-free Tris buffer. In another set of experiments, we used 1.6 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 7.0) buffer in the first step and the Tris buffer in the last step. It was observed that sRNA molecules eluted immediately with 1.6 M $(NH_4)_2SO_4$. These results were important to study the behaviour of the sRNA mixture in different salt concentrations and to develop the best purification strategy for 6S RNA. In fact, high salt concentration plays a key role on the 6S RNA binding to histidine ligand.

Figure 3a shows the chromatographic profile obtained after injection of the sRNA sample (≈100 μg). The presence of different peaks in the chromatogram indicates that the RNAs present in the sRNA population interact differently with the histidine–agarose support. A denaturing polyacrylamide electrophoresis was used to detect and identify the different species eluting in each peak (Figure 3b). The sRNA sample injected on the histidine–agarose matrix (Figure 3b, lane S) was also run on the gel for comparative purposes. This sample was qualitatively characterized by three distinct bands, which corresponded to 6S RNA, 5S RNA and other sRNA, including tRNAs. The electrophoresis analysis indicates that the first peak corresponds to the elution of a small part of 5S RNAs and other sRNA (peak 1 and lane 1 in Figure 3). However, the complete elution of these RNAs was achieved in the second peak, following the reduction of the ionic strength of the elution buffer (peak 2 and lane 2). Finally, with the Tris buffer, 6S RNA was eluted, as it is observed in the third peak of the chromatogram (peak 3 and lane 3).

The main explanation for the specific interactions occurring between the sRNA population and the histidine–agarose matrix is

Copyright © 2010 John Wiley & Sons, Ltd.

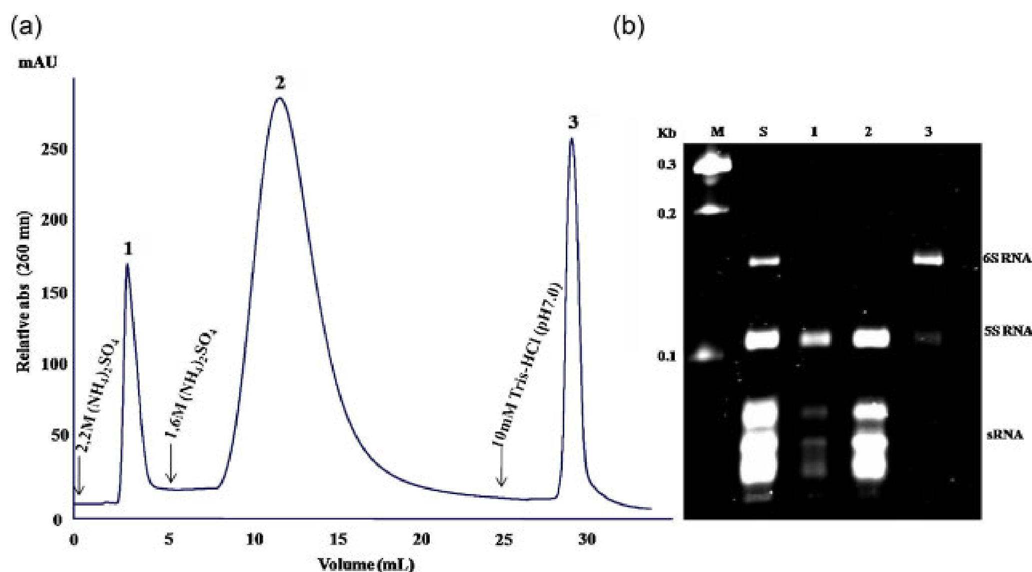

Figure 3. (a) Chromatogram showing the purification of 6S RNA from sRNA population by histidine–agarose chromatography. Step elution was performed at 1 mL/min by stepwise decreasing $(NH_4)_2SO_4$ concentration in the eluent from 2.2 to 1.6 M, finishing with ammonium sulphate-free 10 mM Tris-HCl (pH 7.0) buffer (b) Polyacrylamide gel electrophoresis analysis of the samples collected at the column outlet. Fractions corresponding to peaks (1), (2) and (3) are shown in lanes 1–3, respectively. Lane M, molecular weight marker; lane S, sRNA sample injected onto the column.

the single-stranded nature of RNA, which is normally involved in RNA recognition, due to the high base exposure and availability for interactions. Furthermore, atomic studies performed on protein-RNA complex structures have shown that histidine has a strong tendency to interact with nucleotides (Jeong *et al.*, 2003). As for the exact type of interactions, these may include (i) hydrogen-bonding between H-donor ($N\tau H$) and H-acceptor ($N\pi$) atoms in the non-protonated histidine with base edges; (ii) ring stacking/hydrophobic interactions and (iii) water-mediated hydrogen bonds (Hoffman *et al.*, 2004; Morozova *et al.*, 2006). Thus, and considering the fact that at the working pH (7.0), histidine ($pK_a = 6.5$) is not significantly protonated (Özkara *et al.*, 2002; Pitiot and Vijayalakshmi, 2002) the elution of sRNA when salt concentration is decreased suggests that ring stacking/ hydrophobic interactions and histidine–RNA direct hydrogen-bonding are the dominant effects.

Additionally, 6S RNA structural features seem to be relevant on its distinct behaviour with the histidine–agarose matrix. 6S RNA presents a DNA promoter-like secondary structure consisting of two long irregular double-stranded stem regions, which are interrupted by small bulge loops and a largely single-stranded internal loop in the central region. Along the central bulge and through the continuous stem sequences there are mostly adenines (A) and guanines (G) (Barrick *et al.*, 2005), which were described to interact preferably with histidine (Hoffman *et al.*, 2004). Interestingly, in newly described aspects on the function of 6S RNA, A and G were also identified as specific nucleotides involved in close contact with RNA polymerase (Gildehaus *et al.*, 2007).

Identification of the purified sRNA species

A more accurate identification of 6S RNA species purified by the histidine–agarose matrix was performed. Hence, 6S RNA levels were determined during different stages of *E. coli* growth, taking into account the fact that this RNA is highly expressed in the stationary phase which is already known (Wassarman and Storz, 2000). After the purification process, the fractions of 6S RNA were further identified by reverse-transcription PCR.

The 6S RNA accumulation in the late growing phase of cells is clearly attested in the electrophoretic analysis of the sRNA

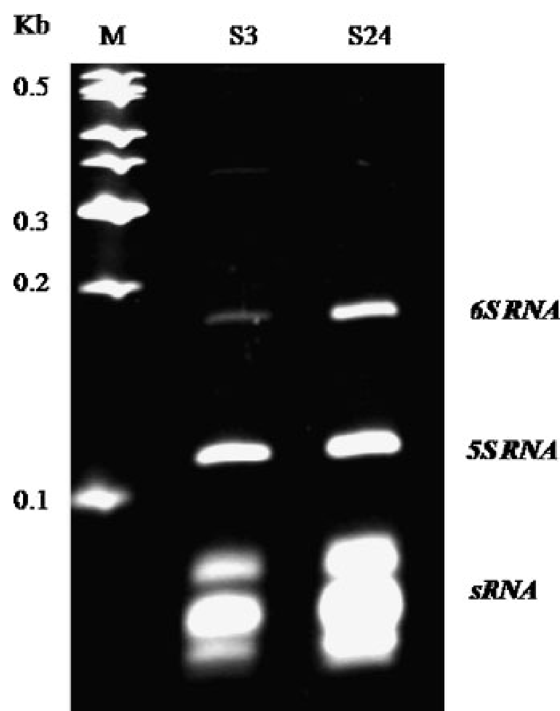

Figure 4. Polyacrylamide electrophoresis analysis of sRNA samples recovered from *E. coli* cell growth was suspended at 3 and 24 h. Cell cultures of *E. coli* DH5α were fermented and their growth was suspended at 3 h and at 24 h of cell growth. Cells were recovered by centrifugation and followed by sRNA extraction procedures. Note the 6S RNA accumulation at 24 h of growth. Lane M, molecular weight marker; Lane S3 and lane S24, sRNA population obtained from 3 and 24 h of cell growth, respectively.

 Copyright © 2010 John Wiley & Sons, Ltd.

samples (Figure 4) because there was a significant increase in 6S RNA concentration from 3 to 24 h of cell growth (Figure 4, lanes S3 and S24). The 6S RNAs from the sRNA populations of the different cell extracts were successfully purified using the above-described chromatographic technique with a histidine–agarose support. For the PCR analysis, the purified 6S RNA pools were treated with DNase I because contaminated DNA could be amplified giving rise to false-positive results. As it can be observed from the electrophoresis analysis of the PCR products (Figure 5), by using specific primers for 6S RNA cDNA, the PCR reaction allowed the amplification of 6S RNA cDNA fragments within 3 h (lane 1) and 24 h (lane 2) of cell growth. No further amplification was seen in the controls (lanes 3–5). The negative control was made of PCR reaction solutions without cDNA, while the positive controls consisted of 6S RNA samples treated with the DNase I, which was used to initially synthesize cDNAs. Therefore, we verified the identity of the sRNA isolated from the original population as the 6S RNA.

The isolation of biologically competent and chemically stable RNA continues to be a central procedure in molecular biology, which has even greater relevance today due to the progress of new disciplines, such as functional genomics and proteomics. The development of new RNA isolation methodologies, such as the purification strategy for 6S RNA isolation that is presented here, is an asset to the research now regularly exploring the transcriptome, the proteome, the metabolome and the genome.

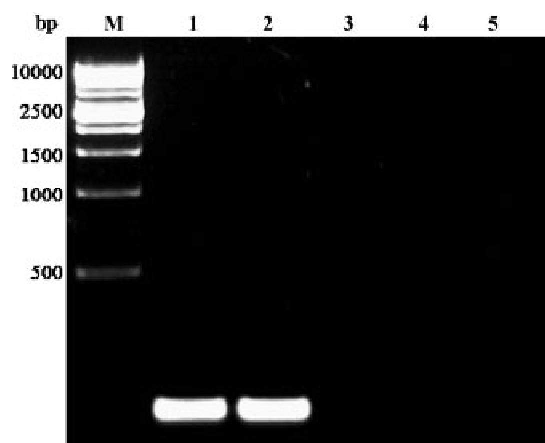

Figure 5. 6S RNA identification by reverse-transcription PCR. The agarose-gel-electrophoretic analysis of PCR products shows amplification of 6S RNA cDNA fragments from *E. coli* cells grown in 3 and 24 h. The controls had no band intensification. Lane M, DNA molecular weight marker; lanes 1 and 2, 6S RNA cDNA from cells cultured for 3 and 24 h, respectively; lane 3, negative control; lanes 4 and 5, 6S RNA samples from cells cultured for 3 and 24 h, respectively, treated with DNase I.

CONCLUSION

In conclusion, this study provided an optimization of the protocol for the separation of sRNA and rRNA from RNAt by reducing the complexity of the general procedure. Moreover, we have developed a purification strategy to separate 6S RNA from *E. coli* sRNA mixture by affinity chromatography with immobilized amino acids, using histidine as a specific ligand. The underlying mechanism involves not only hydrophobic interaction, but also a bio-recognition of nucleic acid bases with histidine. The successful results obtained with this support reveal an efficient technique to obtain a reproducible and appropriate RNA quality with potential applicability for RNA structural and functional studies and gene therapy.

REFERENCES

Barrick JE, Sudarsan N, Weinberg Z, Ruzzo WL, Breaker RR. 2005. 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. *RNA* **11**(5): 774–784.
Chomczynski P, Sacchi N. 2006. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat. Protoc.* **1**(2): 581–585.
Farrell RE. 2005. RNA Methodologies. Academic Press: New York, USA; 688.
Gildehaus N, Neusser T, Wurm R, Wagner R. 2007. Studies on the function of the riboregulator 6S RNA from *E. coli*: RNA polymerase binding, inhibition of *in vitro* transcription and synthesis of RNA-directed *de novo* transcripts. *Nucleic Acids Res.* **35**(6): 1885–1896.
Hoffman MM, Khrapov MA, Cox JC, Yao J, Tong L, Ellington AD. 2004. AANT: the amino acid-nucleotide interaction database. *Nucleic Acids Res.* **32** (Database issue): D174–D181.
Jeong E, Kim H, Lee SW, Han K. 2003. Discovering the interaction propensities of amino acids and nucleotides from protein-RNA complexes. *Mol. Cells* **16**(2): 161–167.
Karen MW. 2007. 6S RNA: a regulator of transcription. *Mol. Microbiol.* **65**(6): 1425–1431.
Morozova N, Allers J, Myers J, Shamoo Y. 2006. Protein-RNA interactions: exploring binding patterns with a three-dimensional superposition analysis of high resolution structures. *Bioinformatics* **22**(22): 2746–2752.
Özkara S, Yavuz H, Patir S, Arica MY, Denizli A. 2002. Separation of human-immunoglobulin-G from human plasma with L-histidine immobilized pseudo-specific bioaffinity adsorbents. *Sep. Sci. Technol.* **37**(3): 717–731.
Pitiot O, Vijayalakshmi MA. 2002. Immobilized histidine ligand affinity chromatography. In Biochromatography: Theory and Practice, Vijayalakshmi MA (ed). 1st ed. Taylor & Francis Inc: New York, USA; 301–325.
Sousa F, Tomaz CT, Prazeres DM, Queiroz JA. 2005. Separation of supercoiled and open circular plasmid DNA isoforms by chromatography with a histidine-agarose support. *Anal. Biochem.* **343**(1): 183–185.
Sousa F, Freitas S, Azzoni AR, Prazeres DM, Queiroz J. 2006. Selective purification of supercoiled plasmid DNA from clarified cell lysates with a single histidine-agarose chromatography step. *Biotechnol. Appl. Biochem.* **45**(Pt 3): 131–140.
Sousa F, Prazeres DM, Queiroz JA. 2008. Affinity chromatography approaches to overcome the challenges of purifying plasmid DNA. *Trends Biotechnol.* **26**(9): 518–525.
Stellrecht CM, Gandhi V. 2002. Concurrent isolation of ribosomal, messenger, and low molecular weight RNA. *Biotechniques* **33**(5): 1122–1124.
Wang YX, Lu M, Draper DE. 1993. Specific ammonium ion requirement for functional ribosomal RNA tertiary structure. *Biochemistry* **32**(46): 12279–12282.
Wassarman KM. 2007. 6S RNA: a small RNA regulator of transcription. *Curr. Opin. Microbiol.* **10**(2): 164–168.
Wassarman KM, Storz G. 2000. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101**(6): 613–623.
Wassarman KM, Zhang A, Storz G. 1999. Small RNAs in *Escherichia coli*. *Trends Microbiol.* **7**(1): 37–45.
Willkomm DK, Hartmann RK. 2005. 6S RNA – an ancient regulator of bacterial RNA polymerase rediscovered. *Biol. Chem.* **386**(12): 1273–1277.

Copyright © 2010 John Wiley & Sons, Ltd.

Paper III

Histidine affinity chromatography-based methodology for the simultaneous isolation of *Escherichia coli* small and ribosomal RNA

R. Martins, J. A. Queiroz, F. Sousa

Biomedical Chromatography. 2012. 26(7): 781-788

Received: 19 July 2011, Accepted: 11 September 2011 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI 10.1002/bmc.1729

Histidine affinity chromatography-based methodology for the simultaneous isolation of *Escherichia coli* small and ribosomal RNA

Rita Martins, João António Queiroz and Fani Sousa*

ABSTRACT: Research on RNA has led to many important biological discoveries and the improvement of therapeutic technologies. In particular, there is a great focus on small RNA and ribosomal RNA owing to their key functions in the cell, which make them excellent therapeutic targets. Although the study of these RNA classes is progressing, some limitations have been found regarding the use of suitable techniques that are able to produce and isolate biologically competent and chemically stable RNA. To address this, we have developed a novel histidine affinity chromatography-based isolation methodology for small and ribosomal RNA molecules. The new procedure involves three main steps: (1) cell lysis with guanidinium buffer, (2) RNA primary isolation with ammonium sulfate precipitation and (3) histidine affinity chromatography to specifically purify small RNA and ribosomal RNA from other *Escherichia coli* impurities (genomic DNA and proteins). The RNA quality assessment revealed that both RNA species were obtained with a high recovery, integrity and purity. The potential of this method to achieve a reproducible RNA isolation with appropriate quality has been demonstrated and it should have broad application in the structural, biophysical and biomedical investigation of systems involving RNA components. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: affinity chromatography; histidine; ribosomal RNA; RNA extraction; small RNA

Introduction

The increasing awareness of the role played by RNA in many biological processes and the progression of diseases makes the discovery of new RNA targets an emerging field for therapeutic treatment. In disciplines such as biotechnology and pharmacology, ribosomal RNA (rRNA) and small RNA (sRNA) are new targets for therapeutic product design and for the development of novel gene-based therapies. RNA is a key drug target for antibacterial treatment, since ribosomal RNA has long been known to be the receptor for antibiotics in clinical practice (Tenson and Mankin, 2006). However, the development of new ribosomal RNA-binding drugs is severely limited by the lack of chemistry studies dedicated to RNA and by the poor understanding of rRNA structure and recognition principles (David-Eden *et al.*, 2010; Foloppe *et al.*, 2006). On the other hand, the noncoding sRNA molecules have been recently recognized as crucial regulatory molecules in all organisms and are revolutionizing medicine owing to their specific functions in the cell, such as gene silencing (Fischer, 2010). The research focusing on these promising molecules is making rapid progress, and sRNA species should continue to be investigated. The possibility of existing undiscovered types of sRNAs in higher organisms (Kawaji *et al.*, 2008; Lagos-Quintana *et al.*, 2001), and the fact that there are still sRNAs whose exact function is unknown, make these molecules a stimulating research area (Gottesman, 2004; Kim, 2005; Wassarman *et al.*, 1999). Therefore, to exploit the potential of rRNA and sRNA, it is necessary to develop new tools to aid in the isolation and purification of large quantities of RNA, as this remains a significant limitation in many structural and biophysical studies (Doudna, 2000).

RNA isolation remains a sensitive and critical process owing to the ubiquitous presence of RNA-degrading enzymes (RNases) in aqueous buffers, on labware, and associated with human handling. The remarkable stability of some forms of RNases contributes to the difficulty of successful handling of RNA in the laboratory. To overcome the problems associated with RNA isolation, several kits are commercially available for safe and easy RNA isolation from a diverse variety of samples. Some of the largest suppliers of such kits are Qiagen (Dusseldorf, Germany), Ambion (Austin, TX, USA) and Promega (Madison, WI, USA) (Chatterjee *et al.*, 2010). The different kits can be used to isolate different types of RNA molecules from different biological sources and provide very similar RNA yield in terms of quality and quantity (Guarino *et al.*, 1997). Most of the existing kits use either silica-based solid-phase extraction (SPE) technology or a magnetic-bead based extraction technology (Boom *et al.*, 1990). Such processes are nondiscriminatory for DNA or RNA, and in most cases the isolation of pure RNA is achieved by a secondary enrichment, either through enzymatic removal of DNA (DNase treatment) or by a second step using specific columns. Also still available are the phenol-based RNA isolation reagents such as TRI reagent (Ambion) or TRIZOL (Invitrogen), which have higher yields than the SPE kits (Chatterjee *et al.*, 2010;

* Correspondence to: F. Sousa, CICS-UBI – Health Sciences Research Centre, University of Beira Interior, Av. Infante D. Henrique, 6200-506 Covilhã, Portugal. E-mail: fani.sousa@fcsaude.ubi.pt

CICS-UBI, Health Sciences Research Centre, University of Beira Interior, Av. Infante D. Henrique 6200-506 Covilhã, Portugal

Abbreviations used: SPE, solid-phase extraction.

Copyright © 2011 John Wiley & Sons, Ltd.

Vomelova *et al.*, 2009). Although phenol-based products are extremely hazardous and the procedure is laborious, this remains the most widely used method in molecular biology to obtain RNA, and is further applied in a large number of analytical techniques (Cristóvão *et al.*, 2009; Elkon *et al.*, 2005; Jin *et al.*, 2009; Maia *et al.*, 2008; Martinho *et al.*, 2010). Nevertheless, none of the existing kits allow the simultaneous isolation of rRNA and sRNA in a single procedure. In addition, very few reports describe methodologies for the concurrent isolation of these RNA classes (Chen *et al.*, 2010; Stellrecht and Gandhi, 2002).

The main goal of this work was to develop a new technique for RNA isolation that enables the isolation and purification of rRNA and sRNA with high integrity and purity, suitable for further applications. To accomplish this purpose, we explored the possibility of using histidine affinity chromatography to selectively isolate the different RNA populations from *Escherichia coli* (*E. coli*) cell lysate. The potential of histidine has already been demonstrated in plasmid DNA purification (Sousa *et al.*, 2006, 2008, 2009b) and recently in RNA purification (Martins *et al.*, 2010). In this last work, we optimized an RNA extraction procedure and developed a new affinity chromatography method using histidine–agarose support to specifically purify 6S RNA from other *E. coli* small RNA species. The study revealed several interesting characteristics of RNA molecules, including their chromatographic behavior and natural interactions that can occur between histidine support and RNA. These results support the interest in applying amino acid-based affinity chromatography to develop new RNA isolation and purification processes.

Materials and method

Material

L-Histidine–agarose gel, the guanidinium salt and all the chemicals used in the lysis buffer were obtained from Sigma-Aldrich (St Louis, MO, USA). Other compounds used in the elution buffer were ammonium sulfate purchased from Panreac (Barcelona, Spain) and Tris from Merck (Darmstadt, Germany). All solutions were freshly prepared using 0.05% diethyl pyrocarbonate (DEPC; Sigma-Aldrich, St Louis, MO, USA)-treated water and the elution buffers were filtered through a 0.20 μm pore size membrane (Schleicher Schuell, Dassel, Germany) and degassed ultrasonically. The DNA molecular weight marker was obtained from Vivantis Technologies (Selangor DE, Malaysia) and RNA ladder was from Invitrogen (Carlsbad, CA, USA). All the materials used in the experiments were RNase-free.

Bacterial growth conditions

The RNA used in this study was obtained from a cell culture of *E. coli* DH5α. Growth was carried out in shake flasks at 37 °C and 250 rpm with 250 mL of Terrific Broth medium (12 g/L tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 0.017 M KH_2PO_4 and 0.072 M K_2HPO_4). Cell growth was suspended in the early exponential phase (OD_{600} = ~2.6). Cells were recovered by centrifugation and stored at −20 °C.

Lysis and RNA primary isolation

Cells were lysed based on a lysis solution reported by Chomczynski and Sacchi (2006). Briefly, 50 mL of bacterial pellets were resuspended by successive pipeting in 5 mL of guanidinium salt buffer [4 M guanidinium thiocyanate; 25 mM sodium citrate, pH 4.0; 0.5% (m/v) *N*-laurosylsarcosine (Sarcosyl); and 0.1 M β-mercaptoethanol) and were then incubated on ice for 10 min. Insoluble material was removed by centrifugation at 19,000 ***g*** for 30 min at 4 °C and the soluble nucleic acids present in the supernatant were concentrated by adding 5 mL of ice-cold isopropanol. The precipitate was recovered by centrifugation at 16,000 ***g*** for 20 min at 4 °C. After centrifuging, the pellet was washed with 2.5 mL of 75% ethanol and incubated at room temperature for 10 min, followed by a 5 min centrifugation at 16,000 ***g*** (4 °C). The air-dried pellet was dissolved in 2 mL of 0.05% DEPC-treated water and incubated for 10 min at 60 °C to ensure complete solubilization. The optical density of the sample was determined at 260/280 nm using a Nanodrop spectrophotometer to assess its quantity and purity. Next, ammonium sulfate was dissolved in the lysate extract up to a final concentration of 1.6 M, followed by 15 min incubation on ice. The precipitated fraction [genomic DNA (gDNA) and rRNA] was separated by centrifugation at 16,000 ***g*** for 30 min at 4 °C, washed with 75% ethanol and solubilized in 0.05% DEPC-treated water. The supernatant (gDNA and sRNA) was desalted and concentrated by Vivaspin concentrators (Vivascience). The concentration and purity of the resulting samples were also determined as previously described. Finally, both samples were loaded directly onto the histidine chromatographic column.

To set a point of reference for the new RNA isolation procedure described, total RNA was also isolated from *E. coli* cells using the commercial solution TRI reagent (Ambion), according to the manufacturer's instructions.

Affinity chromatography

The chromatographic technique was performed in an ÄKTApurifier system with UNICORN software (GE Healthcare, Sweden). A 10 mm diameter × 20 mm long (~2 mL) column was packed with the commercial L-histidine–agarose gel. This support is characterized by the manufacturer as a cross-linked 4% beaded agarose matrix with a 12-atom spacer and an extent of labeling between 1 and 2 μmoL/mL. The experiments performed with the partially purified sRNA molecules were initiated with the equilibration of the column with 2.0 M of $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 8.0) buffer at a flow-rate of 1 mL/min while the experiments for rRNA purification were performed equilibrating the column with 1.5 M of $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 8.0). In the two set of experiments, both samples were injected onto the column using a 100 μL loop at the same flow-rate. The absorbance of the eluates was continuously monitored at 260 nm. After the elution of unbound species with 2.0 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 8.0) buffer or 1.5 M $(NH_4)_2SO_4$ in 10 mM Tris–HCl (pH 8.0), the ionic strength of the buffer was decreased to ammonium sulfate-free 10 mM Tris–HCl (pH 8.0) buffer. Chromatographic runs were performed at room temperature (~25 °C). The fractions were pooled according to the chromatograms obtained, and following concentration and desalting with Vivaspin concentrators, the pools were kept for quantification and further analysis.

Agarose gel electrophoresis

Pooled fractions were analyzed by horizontal electrophoresis using 15 cm 1.2% agarose gels (Hoefer, San Francisco, CA, USA), stained with ethidium bromide (0.5 μg/mL). Electrophoresis was carried out at 120 V with TAE buffer in DEPC-treated water

Copyright © 2011 John Wiley & Sons, Ltd.

(40 mM Tris base, 20 mM acetic acid and 1 mM EDTA, pH 8.0). The nucleic acids were visualized using a Vilber Lourmat system (ILC Lda, Lisbon, Portugal).

Polyacrylamide electrophoresis

The sRNA supernatant and the fractions recovered from the sRNA chromatographic experiment were also analyzed by vertical electrophoresis using an Amersham Biosciences system (GE Healthcare, Sweden) with 10% polyacrylamide gel. Electrophoresis was carried out at 120 V for 90 min with TBE buffer (0.84 M Tris base, 0.89 M boric acid and 0.01 M EDTA, pH 8.3). sRNA samples were previously denatured with 97.5% formamide and denatured conditions were kept in the gel owing to the presence of 8 M urea. Thus, sRNA molecules in the gel were visualized using the Vilber Lourmat system after staining with ethidium bromide (0.5 μg/mL).

Protein analysis

The protein concentration was measured by the micro-BCA (bicinchoninic acid) assay (Thermo Fisher Scientific Inc.). A fraction of each sample (50 μL) was added to 200 μL of BCA reagent in a microplate and incubated for 30 min at 60 °C. Absorbance was measured at 570 nm in a microplate reader. The calibration curve was prepared using bovine serum albumin standards (0.01–0.1 mg/mL).

Results and discussion

The importance of RNA analysis has increased owing to the progress of basic and applied science considering this macromolecule, as well as its application in clinical fields. Although rRNA and sRNA are two important classes of RNA that can trigger great therapeutic evolution, the lack of structural and chemical information on these molecules delays the enhancement of medical strategies for severe diseases. Therefore, the development of new tools for the isolation of RNA species provides greater availability of pure molecules that certainly will aid studies.

This work describes a new method in which rRNA and sRNA can be isolated from *E. coli* cell lysate, demonstrating the successful application of amino acid-based affinity chromatography using histidine as a biospecific ligand in the purification of RNA species from gDNA and proteins.

Figure 1 schematically shows the new chromatographic-based method for the isolation and purification of both RNA classes. Each step of the method is detailed and discussed in the following sections. RNA was also isolated from *E. coli* cells with TRI reagent from Ambion, a commonly used commercial reagent based on phenol–chloroform methodology, to compare the efficacy and applicability.

E. coli cell lysis

The first steps of the new RNA isolation procedure include *E. coli* cell lysis with guanidinium-based lysis buffer and the removal of cell debris to obtain a clarified cell lysate (Fig. 1 A). Guanidinium thiocyanate is a stronger protein denaturant that appears repeatedly in the literature and in various commercial formulations for the preparation of RNA from sources enriched in RNase activity (Farrell, 2005). In general, the inclusion of guanidinium buffers in the majority of the procedures relies on their: (1) efficiency in protein denaturation, inhibiting RNases activity; (2) recovery of high-quality RNA samples; and (3) chaotropicity. Figure 2, lane L, shows the nucleic acid profile of the clarified lysate, containing different RNA species and gDNA obtained after cell lysis. The total RNA sample obtained with TRI reagent (Fig. 2, lane T) was also loaded on the gel for comparative purposes. It is important to note that, unlike the samples resulting from TRI extraction that only contain RNA species, the presence of gDNA in the lysate was expected, since, at this stage of the new method, no further purification has been performed. Nevertheless, the electrophoresis indicates an accurate banding profile for both samples compared with the manufacturing information of the commercial reagent. As observed in the electrophoresis of the lysate (Fig. 2, lane L), the use of guanidinium-based lysis buffers mandates a procedure for the partitioning of both RNA classes and DNA. Therefore, we have developed a new strategy to accomplish the successful isolation and purification of sRNA and rRNA, combining salt precipitation (Fig. 1 B) and affinity-chromatography (Fig. 1 C).

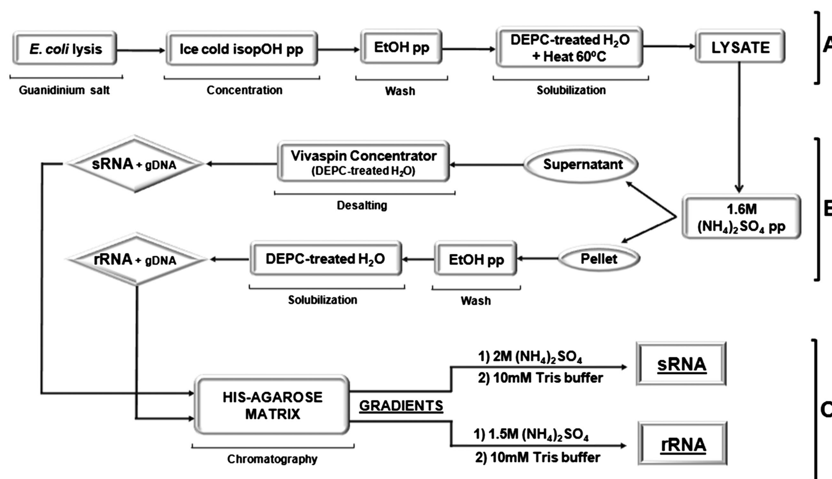

Figure 1. Overall procedure of the new methodology developed for isolation of RNA species. The method includes *E. coli* cell lysis (A) and an ammonium sulfate precipitation step to get an initial separation between the different RNA populations (B). Finally, RNA purification is accomplished by histidine affinity chromatography (C). (DEPC-treated H_2O, diethyl pyrocarbonate-treated water; etOH, ethanol; His, histidine; IsopOH, isopropanol; pp, precipitation).

Copyright © 2011 John Wiley & Sons, Ltd.

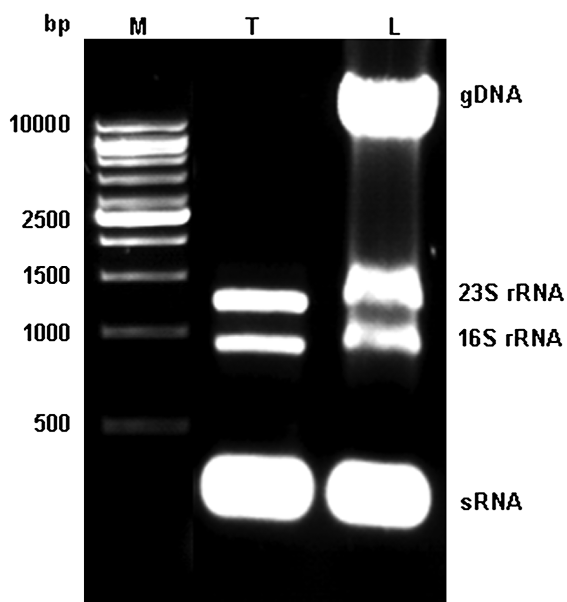

Figure 2. Agarose gel electrophoresis analysis of *E. coli* cell lysate recovered by guanidinium-based lysis and total RNA isolated with TRI reagent. Lane M, DNA molecular weight marker; lane T, total RNA sample isolated with TRI reagent; lane L, cell lysate obtained by primary isolation.

sRNA and rRNA primary isolation

In order to accomplish pure RNA populations from the cell lysate, sRNA and rRNA were first separated from each other by a precipitation step (Fig. 1 B). In a previous work (Martins *et al.*, 2010), we have successfully exploited RNA stringency properties, which enables rRNA separation from sRNA population through an ammonium sulfate precipitation. Building on these results, the separation between sRNA and rRNA in the crude lysate was accomplished by adjusting the ionic strength of the lysate solution with 1.6 M of $(NH_4)_2SO_4$, as schematically shown in Fig. 1(B). The nucleic acid content of the supernatant and the pellet samples was analyzed by agarose gel electrophoresis as shown in Fig. 3. The sRNA and gDNA were recovered in the supernatant (Fig. 3, lane S) while the pellet contained rRNA and a few gDNA molecules (Fig. 3, lane P). Thus, the ammonium sulfate precipitation was revealed to be an interesting approach for the preliminary separation of RNA molecules from cell lysates.

sRNA and rRNA chromatographic purification

Affinity chromatography is a unique separation method, as it is the only technique that permits the purification of biomolecules on the basis of biological functions rather than individual physical or chemical properties. Therefore, the affinity chromatography has been widely used in downstream processes, owing to the highly specific interaction between the ligand and target biomolecules. Our research group has been developing a systematic and continuous work in plasmid DNA purification highlighting the potential of amino acid-based affinity chromatography for this purpose (Sousa *et al.*, 2005, 2006, 2008, 2009a, 2009c, 2010; Sousa and Queiroz, 2011). These studies revealed the possibility to develop new purification approaches by exploiting favored interactions of amino acids and other nucleic acids. In fact, RNA purification using amino-acid based affinity chromatography was recently initiated with an isolation strategy for *E. coli* 6S RNA using histidine affinity chromatography (Martins *et al.*, 2010). In that work, the histidine support was able to isolate 6S RNA from a mixture of sRNAs owing to a biorecognition phenomena occurring between the ligands and the RNA bases.

The present work evaluates the possibility of exploiting favored interactions between histidine and the different nucleic acids present in an *E. coli* lysate in order to develop an affinity chromatographic operation to specifically purify each RNA population from DNA and proteins (Fig. 1 C), from a complex cell lysate.

Different chromatographic conditions were required to selectively purify sRNA and rRNA molecules. In order to evaluate and adjust the chromatographic conditions for purification purposes, a set of chromatographic assays were first performed to study the ionic strength effects on RNA retention (data not shown). Briefly, $(NH_4)_2SO_4$ concentrations of 2 and 1.6 M were tested for impure sRNA samples. These preliminary studies showed that the higher 2 M salt concentration promotes total sRNA retention but gDNA is readily eluted, while the application of 1.6 M $(NH_4)_2SO_4$ to the histidine column causes the immediate elution of both nucleic acids. The behavior of sRNA molecules in these experiments is in accordance with the results previously described (Martins *et al.*, 2010), where the retention pattern of sRNA population was evaluated to accomplish 6S RNA isolation. In this previous work, the sRNA species were totally retained with 2.2 M $(NH_4)_2SO_4$, and its elution was achieved with the reduction of the ionic strength of the elution buffer to 1.6 M $(NH_4)_2SO_4$. In the present work, the main goal is the isolation of the whole sRNA population from a mixture containing gDNA molecules, thus the specificity of histidine–agarose matrix, which selectively retains sRNA molecules even in more complex feed samples, is noteworthy.

Because ionic strength influences RNA conformational chemistry (Farrell, 2005), particularly in rRNA species (Martins *et al.*, 2010), the concentrations of $(NH_4)_2SO_4$ tested for rRNA samples were up to 1.5 M, which also prevents precipitation. The results demonstrated that below 1.5 M $(NH_4)_2SO_4$, gDNA molecules are co-eluted with rRNA, which precludes accurate rRNA purification. Thus, the best condition to accomplish rRNA isolation from gDNA is the application of 1.5 M $(NH_4)_2SO_4$.

Based on the above preliminary studies, it was possible to develop a new purification strategy using the histidine–agarose support, to successfully attain isolated and pure sRNA and rRNA molecules. Figure 4 shows the chromatograms with the

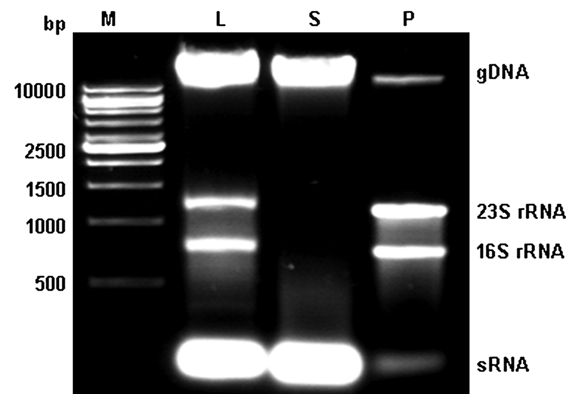

Figure 3. Agarose gel electrophoresis analysis of the *E. coli* cell lysate after ammonium sulfate precipitation. Lane M, DNA molecular weight marker; lane L, cell lysate; lane S, supernatant of salt precipitation; Lane P, pellet formed in the precipitation step.

Copyright © 2011 John Wiley & Sons, Ltd.

optimized ionic strength conditions applied to histidine–agarose chromatography for the purification of each RNA population, sRNA and rRNA (Fig. 4A and B, respectively) and the electrophoresis analysis of the pure factions (Fig. 4C). The presence of two main peaks in both chromatograms indicates that the nucleic acids present in the samples interact differently with the histidine support. The agarose gel electrophoresis was used to detect and identify the nucleic acids eluting in each peak (Fig. 4C). The impure sRNA and rRNA samples injected (Fig. 4, lanes S and P, respectively) were also run on the gel for comparative purposes. The electrophoretic results indicate that the first peak obtained in both chromatograms (Fig. 4A, peak 1 and B, peak 3) corresponds to gDNA (Fig. 4C, lanes 1 and 3). This nucleic acid was not retained in the histidine–agarose support at either 2.0 or 1.5 M $(NH_4)_2SO_4$ salt concentration. Curiously, in the rRNA purification chromatogram were observed two first peaks (Fig. 4B, peaks 3) with similar retention, both identified as gDNA by electrophoresis analysis (Fig. 4C, lane 3). In addition, gDNA was also verified both in supernatant and pellet samples after the ammonium sulfate precipitation step (Fig. 3). These two facts suggest the existence of a heterogeneous mixture of gDNA molecules, probably owing to the lysis process. The possibility of the existence of single- and double-stranded gDNA molecules, which provide size heterogeneity in the lysate solution, explains the ambiguous behavior of gDNA during the RNA isolation method.

To recover the bound sRNA and rRNA from the histidine–agarose matrix, the elution buffer was changed to ammonium sulfate-free 10 mM Tris–HCl (pH 8) buffer, as can be observed from the electrophoresis analysis of the peaks 2 and 4 (Fig. 4C, lanes 2 and 4, respectively), where both sRNA and rRNA were recovered with high purity.

The different interaction behavior of double- and single-stranded nucleic acids during the purification process has been described (Sousa *et al*., 2006, 2009b), where plasmid and gDNA presented lower retention when compared with RNA and single-stranded oligonucleotides. In line with these results, gDNA was observed to be less retained, revealing a weak affinity for the histidine–agarose support. This could be due to the double stranded structure of this biomolecule, which causes the coverage of the principal contact surfaces, the nucleotide bases (Sousa *et al*., 2006). On the other hand, the stronger and selective interaction of both isolated RNA species with the histidine–agarose matrix can be explained by the single-stranded structure of RNA, which is normally involved in RNA recognition, owing to the high nucleotide base exposure and availability for interactions. Furthermore, atomic studies performed on protein–RNA complex structures have shown that histidine has a strong tendency to interact with nucleotides (Jeong *et al*., 2003). As for the exact type of interactions, these may include (i) hydrogen-bonding between H-donor ($N\tau H$) and H-acceptor ($N\pi$) atoms in the nonprotonated histidine with base edges; (ii) ring stacking/hydrophobic interactions; and (iii) water-mediated hydrogen bonds (Hoffman *et al*., 2004; Morozova *et al*., 2006). Moreover, analysis of nucleotides binding to histidine chromatography supports were recently provided by STD-NMR spectroscopy (Cruz *et al*., 2011). This technique enabled the binding characterization and understanding of the interactions involved in the recognition of 5′-mononucleotides by the histidine support. Thus, and considering the fact that, at the working pH (8.0), histidine ($pK_a = 6.5$) is not significantly protonated (Özkara *et al*., 2002; Pitiot and Vijayalakshmi, 2002) the elution of RNA when salt concentration is decreased suggests that ring stacking/

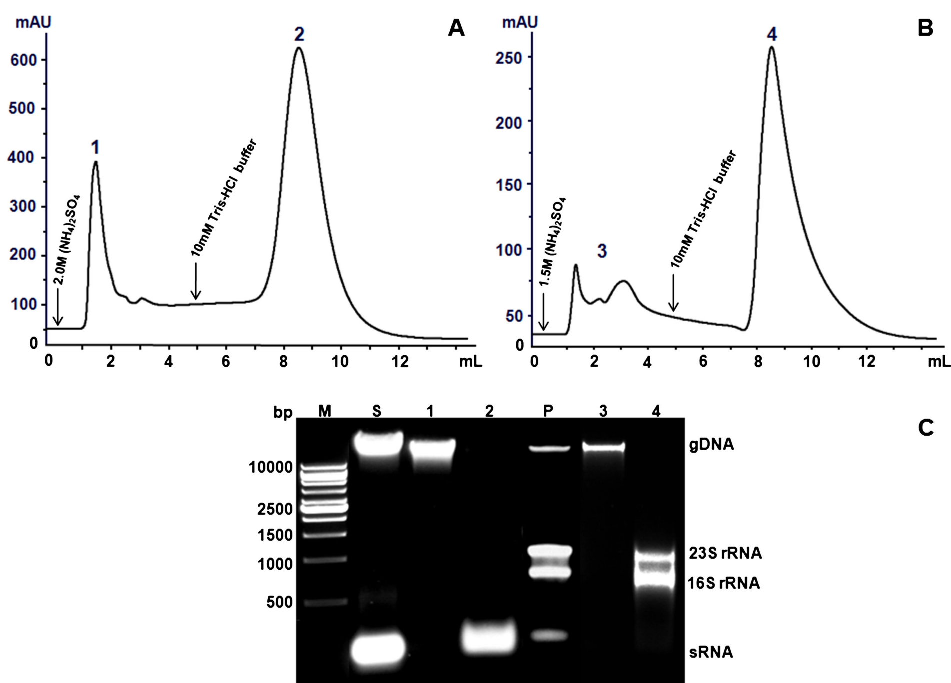

Figure 4. Affinity chromatographic profile of nucleic acid retention on histidine–agarose support, at 25 °C. (A) sRNA purification chromatogram. (B) rRNA purification chromatogram. (C) Agarose gel electrophoresis analysis of the samples collected at the column outlet. Fractions corresponding to peaks **1–4** are shown in lanes 1–4, respectively. Lane M, DNA molecular weight marker; lane S, sample obtained from the supernatant of salt precipitation injected onto the column; lanes 1 and 3, gDNA; lane 2, sRNA molecules; lane P, sample obtained from the pellet formed in the precipitation step injected onto the column; lane 4, rRNA molecules 16S and 23S.

Copyright © 2011 John Wiley & Sons, Ltd.

Table 1. Quantitative analysis of the cell lysate throughout the primary isolation process. Data are presented as means with SD ($n = 3$)

Sample	Volume (mL)	Nucleic acids (μg/mL)	Nucleic acids (μg)	Ratio (260/280 nm)	Proteins (μg/mL)	Proteins (μg)
Cell lysate	2.00	1132 ± 12.77	2264.00 ± 25.53	1.97 ± 0.02	206.15 ± 5.32	412.30 ± 10.65
$(NH_4)_2SO_4$ precipitation						
Supernatant	2.00	890.00 ± 16.93	1780 ± 33.86	1.96 ± 0.01	53.08 ± 4.88	106.15 ± 9.75
Pellet	0.50	478.00 ± 6.35	239 ± 3.18	2.04 ± 0.01	140.28 ± 7.03	70.14 ± 3.52

hydrophobic interactions and histidine–RNA direct hydrogen-bonding are the dominant effects.

The present data shows the development of an affinity chromatography approach based on natural occurrence interactions, between histidine and nucleic acids, aiming the purification of RNA molecules. The exploitation of these affinity interactions can trigger new insights not only in isolation strategies but also in many other RNA research fields owing to its implication in molecular recognition phenomena.

Assessment of sRNA and rRNA yield and quality

To ensure the success of the purification methodology here described, it is essential to control the quality of RNA preparations. In order to check RNA quality, two strategies are commonly followed: spectrophotometry analysis and agarose gel electrophoresis (Farrell, 2005; Madabusi *et al.*, 2006). The RNA recovery yield and purity were determined by measuring absorbance at 260 and 280 nm. The spectrophotometry analysis estimated the nucleic acid concentration using the absorbance values at 260 nm, while the purity of each sample was determined by calculating the 260/280 ratio. Electrophoresis analysis is an important complementary technique that enables assessing RNA integrity and purity, as well.

The quantitative analysis of purity and recovery yield of RNA throughout the isolation procedure is summarized in Tables 1 and 2. In sRNA and rRNA primary separation (Table 1), the spectrophotometric data show that the ammonium sulfate precipitation step has a good recovery yield, because 1780 ± 33.86 μg of nucleic acids was recovered in the supernatant and 239 ± 3.18 μg in the pellet, considering the initial mass of nucleic acids in the cell lysate (2264 ± 25.53 μg). In addition, the ratio between the absorbance at 260 nm and 280 nm for lysate samples was below the reference ratio of 2.00 ± 0.1 for a pure RNA sample (Farrell, 2005). The low ratio usually indicates protein contamination owing to protein absorption at 280 nm. Hence, total protein content was evaluated by micro-BCA assay. The results revealed that protein concentration is practically maintained during the precipitation process. However, proteins tend to precipitate along with rRNA. The rRNA pellet presented a concentration of 140.28 ± 7.03 μg/mL of protein, which is almost 3-fold the protein measured in the supernatant (53.08 ± 4.88 μg/mL).

Table 2 shows the results for purity and recovery yield of RNA species purified by histidine affinity chromatography. The chromatographic process demonstrates an excellent performance in the isolation of each RNA species, attaining a recovery yield of 90.0% for sRNA and 51.5% for rRNA molecules. To establish a point of reference for this new RNA method, total RNA was isolated with TRI reagent. TRI samples were obtained with 66.00 ± 0.24 μg of total RNA while the total mass of RNA species recovered by the new RNA purification strategy was 69.6 ± 0.67 μg (56.00 ± 0.24 μg of sRNA and 13.60 ± 0.69 μg of rRNA; Table 2). Therefore, the overall yield of the proposed RNA isolation technique is comparable to that of phenol-based methodologies, which have been described as the most effective methods in RNA isolation in terms of yield and quality (Sambrook *et al.*, 2001).

The purity of the samples was also evaluated by measuring the 260/280 nm ratio. Similar to total RNA sample isolated with TRI reagent, sRNA and rRNA species have shown 260/280 nm ratios of 2.00 ± 0.01, which is often characteristic of a pure RNA preparation (Farrell, 2005). The protein content is a common impurity in RNA preparations that can be omitted in spectrophotometry analysis. In fact, a relatively large amount of protein contamination is necessary to significantly affect the 260/280 ratio

Table 2. Quantitative analysis of purity and recovery yield of the RNA species isolated by histidine–agarose chromatography. Data are presented as means with SD ($n = 3$)

Sample	Nucleic acids (μg) gDNA	Nucleic acids (μg) RNA	Recovery yield (%)	Ratio (260/280 nm)	Proteins (μg/mL)	Proteins (μg)
Supernatant	89.00 ± 1.71			1.96 ± 0.02	53.08 ± 4.88	5.31 ± 0.49
Peak 1	23.00 ± 1.20	ND	90.00	2.00 ± 0.02	22.82 ± 7.10	2.28 ± 0.71
Peak 2	ND	56.00 ± 0.26		2.00 ± 0.00	30.67 ± 8.11	3.07 ± 0.81
Pellet	47.80 ± 0.64			2.04 ± 0.01	140.28 ± 7.03	70.14 ± 3.52
Peak 3	11.00 ± 0.30	ND		1.86 ± 0.02	15.08 ± 1.52	1.51 ± 0.15
Peak 4	ND	13.60 ± 0.69	51.50	2.15 ± 0.01	22.49 ± 7.04	2.25 ± 0.70
TRI reagent[a]	ND	66.00 ± 0.24	—	2.00 ± 0.01	41.35 ± 0.16	4.14 ± 0.02

ND, not detected by agarose gel electrophoresis.
[a]Control sample.

Copyright © 2011 John Wiley & Sons, Ltd.

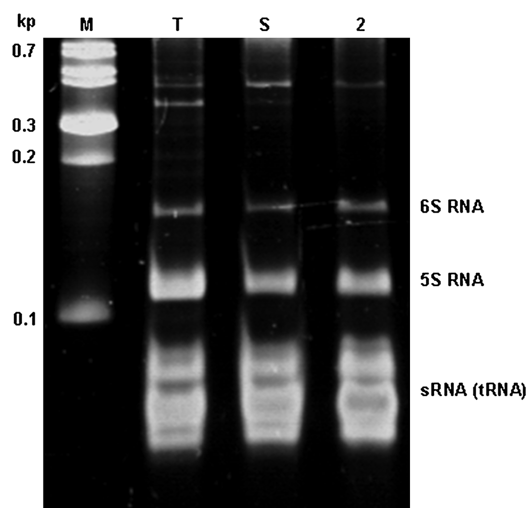

Figure 5. Polyacrylamide gel electrophoresis analysis of sRNA molecules. Lane M, RNA molecular weight marker; lane T, sRNA population of total RNA isolated with TRI reagent; lane S, sample of the supernatant of salt precipitation injected onto the column; lane 2, sRNA molecules purified by histidine–agarose support.

in an RNA solution (Sambrook *et al*., 2001). In addition, the protein analysis in the sample of total RNA isolated with TRI reagent indicates the presence of proteins in a concentration of $41.35 \pm 0.16\,\mu g/mL$, which suggests that RNA samples are not entirely pure, even presenting a ratio of 2.00. Thus, protein content was also measured in the RNA fractions collected from the chromatographic peaks. It was found that the chromatographic gradient used to purify sRNA eliminates almost half of the proteins ($22.82 \pm 7.10\,\mu g/mL$) contained in the supernatant sample throughout the simultaneous elution of proteins with gDNA at high salt concentration [2 M $(NH_4)_2SO_4$]. As a result, the protein contamination level was significantly reduced in sRNA samples ($30.67 \pm 8.11\,\mu g/mL$). Curiously, in the rRNA gradient the reduction of proteins is more pronounced as only $22.49 \pm 7.04\,\mu g/mL$ is present in the final rRNA sample. The analysis of protein content in the fractions collected along the rRNA purification gradient revealed that proteins tend to elute steadily throughout the gradient (data not show). Therefore, the chromatographic strategy used in this new RNA isolation method also provides a reduction of protein contamination in both sRNA and rRNA samples, achieving a similar protein concentration as in total RNA samples isolated with TRI reagent.

The previous analyses allowed the assessment of RNA quality, indicating that the developed RNA isolation methodology can purify the different RNA species with high recovery yield and purity. However, electrophoresis analysis allows the visualization of the nucleic acid components of a sample, which is crucial to guarantee RNA quality. The integrity of the rRNA subunits (23S, 16S and 5S), the presence of low molecular weight RNA-degradation products and the presence of gDNA contamination are commonly assessed using agarose gel electrophoresis. Ideally, all expected rRNA subunits should be observed, with no signs of RNA degradation products or presence of gDNA (Farrell, 2005). Figure 4(C) shows a native agarose gel electrophoresis analysis performed for both purified RNA populations. Similarly to what was achieved for the commercial RNA preparation (Fig. 2), RNA bands are sharp and clear with a characteristic banding profile. These results are good indicators of chemically intact and biologically competent sRNA and rRNA (Fig. 4C, lanes 2 and 4). In addition, the polyacrylamide gel electrophoresis analysis in Fig. 5 confirmed that there was no degradation of sRNA population and the integrity of 5S RNA was maintained. These conclusions can be assumed owing to the similar electrophoretic profile for sRNA isolated with TRI reagent (lane T) and with the new RNA isolation strategy (lanes S and 2).

The successful isolation of sRNA and rRNA species with high yield in terms of quality and quantity is a step forward in RNA extraction methodologies and can widely contribute to studies in many RNA research fields. Affinity chromatography is a unique technique that simulates and exploits biological processes such as molecular recognition for the selective purification of the target molecule, while eliminating additional steps, increasing yields and improving process economics. In a single step, histidine affinity purification offered RNA isolation from a complex biological mixture, eliminating the use of other chemicals during the isolation process. The chromatographic technique simplifies workflow integration and miniaturizes sample handling process, making it useful for the development of an RNase and organic solvents free methodology. Moreover, an RNA isolation procedure based on affinity columns can be automated by the development of specialized equipment. Over recent years, demand for automated systems designed for medium to large laboratories has grown. Automating nucleic acid extraction processes is potentially beneficial for a number of reasons, including reducing working time, decreasing labor costs and increasing worker safety. At the same time they provide opportunities to increase reproducibility and quality of results (Chatterjee *et al*., 2010; Tan and Yiap, 2009).

Conclusion

The present work demonstrates the development of a novel protocol for isolation and purification of different RNA classes. The peculiar chemical properties of RNA allowed the separation of two main RNA populations, sRNA and rRNA, from *E. coli* cell lysates, whereas histidine affinity chromatography enables their selective purification, eliminating gDNA and proteins. The chromatographic system with histidine–agarose is demonstrated to be a multifunctional technique, which attests to its successful performance in RNA purification, supported by accurate recovery yields and purity degree achieved with this new method. To conclude, the new isolation methodology for sRNA and rRNA represents a major advance in the ability to accurately purify RNA molecules that can also be applicable to a broad range of biochemical applications.

Acknowledgments

This work was supported by FCT, the Portuguese Foundation for Science and Technology (PTDC/EBB-BIO/114320/2009). Rita Martins also acknowledges a fellowship (SFRH/BD/ 64100/2009) from FCT.

References

Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM and van der Noordaa J. Rapid and simple method for purification of nucleic acids. *Journal of Clinical Microbiology* 1990; **28**(3): 495–503.

Chatterjee A, Mirer PL, Zaldivar Santamaria E, Klapperich C, Sharon A and Sauer-Budge AF. RNA isolation from mammalian cells using porous

Copyright © 2011 John Wiley & Sons, Ltd.

polymer monoliths: an approach for high-throughput automation. *Analytical Chemistry* 2010; **82**(11): 4344–4356.

Chen P, Jager G and Zheng B. Transfer RNA modifications and genes for modifying enzymes in *Arabidopsis thaliana*. *BMC Plant Biology* 2010; **10**(1): 201.

Chomczynski P and Sacchi N. The single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction: twenty-something years on. *Nature Protocols* 2006; **1**: 581–585.

Cristóvão AC, Choi D-H, Baltazar G, Beal MF and Kim Y-S. The role of NADPH oxidase 1-derived reactive oxygen species in paraquat-mediated dopaminergic cell death. *Antioxidants & Redox Signaling* 2009; **11**(9): 2105–2118.

Cruz C, Cabrita EJ and Queiroz JA. Analysis of nucleotides binding to chromatography supports provided by nuclear magnetic resonance spectroscopy. *Journal of Chromatography. A* 2011; **1218**(22): 3559–3564.

David-Eden H, Mankin AS and Mandel-Gutfreund Y. Structural signatures of antibiotic binding sites on the ribosome. *Nucleic Acids Research* 2010; **38**(18): 5982–5994.

Doudna JA. Structural genomics of RNA. *Nature Structural Biology* 2000; **7**(suppl.): 954–956.

Elkon R, Rashi-Elkeles S, Lerenthal Y, Linhart C, Tenne T, Amariglio N, Rechavi G, Shamir R and Shiloh Y. Dissection of a DNA-damage-induced transcriptional network using a combination of microarrays, RNA interference and computational promoter analysis. *Genome Biology* 2005; **6**(5): R43.

Farrell RE. *RNA Methodologies*, 3rd edn. Academic Press: New York, 2005.

Fischer SE. Small RNA-mediated gene silencing pathways in *C. elegans*. *The International Journal of Biochemistry & Cell Biology* 2010; **42**(8): 1306–1315.

Foloppe N, Matassova N and Aboul-ela F. Towards the discovery of drug-like RNA ligands? *Drug Discovery Today* 2006; **11**(21–22): 1019–1027.

Gottesman S. The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annual Review of Microbiology* 2004; **58**: 303–328.

Guarino H, Goyal SM and Murtaugh MP. Comparison of four RNA extraction methods for the detection of porcine reproductive and respiratory syndrome virus by RT-PCR. *The New Microbiologica* 1997; **20**: 319–324.

Hoffman MM, Khrapov MA, Cox JC, Yao J, Tong L and Ellington AD. AANT: the Amino Acid-Nucleotide Interaction Database. *Nucleic Acids Research* 2004; **32**: 174–181.

Jeong E, Kim H, Lee SW and Han K. Discovering the interaction propensities of amino acids and nucleotides from protein–RNA complexes. *Molecules and Cells* 2003; **16**: 161–167.

Jin ZB, Hirokawa G, Gui L, Takahashi R, Osakada F, Hiura Y, Takahashi M, Yasuhara O and Iwai N. Targeted deletion of miR-182, an abundant retinal microRNA. *Molecular Vision* 2009; **15**: 523–533.

Kawaji H, Nakamura M, Takahashi Y, Sandelin A, Katayama S, Fukuda S, Daub C, Kai C, Kawai J, Yasuda J, Carninci P and Hayashizaki Y. Hidden layers of human small RNAs. *BMC Genomics* 2008; **9**(1): 157.

Kim VN. Small RNAs: classification, biogenesis, and function. *Molecules and Cells* 2005; **19**(1): 1–15.

Lagos-Quintana M, Rauhut R, Lendeckel W and Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science* 2001; **294**(5543): 853–858.

Madabusi LV, Latham GJ and Andruss BF. RNA extraction for arrays. In *Methods in Enzymology*, Alan K and Brian O (eds). Academic Press: San Diego, CA, 2006; 1–14.

Maia C, Socorro S, Schmitt F and Santos C. STEAP1 is over-expressed in breast cancer and down-regulated by 17β-estradiol in MCF-7 cells and in the rat mammary gland. *Endocrine* 2008; **34**(1): 108–116.

Martinho A, Gonçalves I, Cardoso I, Almeida MR, Quintela T, Saraiva MJ and Santos CRA. Human metallothioneins 2 and 3 differentially affect amyloid-beta binding by transthyretin. *FEBS Journal* 2010; **277**(16): 3427–3436.

Martins R, Queiroz JA and Sousa F. A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography. *Journal of Molecular Recognition* 2010; **23**(6): 519–524.

Morozova N, Allers J, Myers J and Shamoo Y. Protein–RNA interactions: exploring binding patterns with a three-dimensional superposition analysis of high resolution structures. *Bioinformatics* 2006; **22**: 2746–2752.

Özkara S, Yavuz H, Patr S, Arca MY and Denizli A. Separation of human-immunoglobulin-G from human plasma with L-histidine immobilized pseudo-specific bioaffinity adsorbents. *Separation Science and Technolology* 2002; **37**: 717–731.

Pitiot O and Vijayalakshmi MA, Immobilized histidine ligand affinity chromatography. In *Biochromatography: Theory and Practice*, Vijayalakshmi MA (ed.). Taylor & Francis: New York, 2002; 301–324.

Sambrook J, Fritsch EF and Maniatis T. *Molecular Cloning: a Laboratory Manual, 3rd edn*. Cold Spring Harbor Laboratory Press: New York, 2001.

Sousa F and Queiroz JA. Supercoiled plasmid quality assessment by analytical arginine-affinity chromatography. *Journal of Chromatography. A* 2011; **1218**(1): 124–129.

Sousa F, Tomaz CT, Prazeres DM and Queiroz JA. Separation of supercoiled and open circular plasmid DNA isoforms by chromatography with a histidine–agarose support. *Analytical Biochemistry* 2005; **343**: 183–185.

Sousa F, Freitas S, Azzoni AR, Prazeres DM and Queiroz JA. Selective purification of supercoiled plasmid DNA from clarified cell lysates with a single histidine–agarose chromatography step. *Biotechnology and Applied Biochemistry* 2006; **45**: 131–140.

Sousa F, Prazeres DM and Queiroz JA. Affinity chromatography approaches to overcome the challenges of purifying plasmid DNA. *Trends in Biotechnology* 2008; **26**(9): 518–525.

Sousa A, Sousa F and Queiroz JA. Biorecognition of supercoiled plasmid DNA isoform in lysine-affinity chromatography. *Journal of Chromatography B* 2009a; **877**(27): 3257–3260.

Sousa A, Sousa F, Prazeres DM and Queiroz JA. Histidine affinity chromatography of homo-oligonucleotides. Role of multiple interactions on retention. *Biomedical Chromatography* 2009b; **23**: 745–753.

Sousa A, Sousa F and Queiroz JA. Differential interactions of plasmid DNA, RNA and genomic DNA with amino acid-based affinity matrices. *Journal of Separation Science* 2010; **33**(17–18): 2610–2618.

Sousa F, Prazeres DMF and Queiroz JA. Improvement of transfection efficiency by using supercoiled plasmid DNA purified with arginine affinity chromatography. *The Journal of Gene Medicine* 2009c; **11**(1): 79–88.

Stellrecht CM and Gandhi V. Concurrent isolation of ribosomal, messenger, and low molecular weight RNA. *Biotechniques* 2002; **33**(5): 1122–1124.

Tan SC and Yiap BC. DNA, RNA, and protein extraction: the past and the present. *Journal of Biomedicine and Biotechnology* 2009; **2009**: 1–10.

Tenson T and Mankin A. Antibiotics and the ribosome. *Molecular Microbiology* 2006; **59**(6): 1664–1677.

Vomelova I, Vanickova Z and Sedo A. Methods of RNA purification. All ways (should) lead to Rome. *Folia Biologica (Praha)* 2009; **55**(6): 243–251.

Wassarman KM, Zhang A and Storz G. Small RNAs in *Escherichia coli*. *Trends in Microbiology* 1999; **7**(1): 37–45.

Copyright © 2011 John Wiley & Sons, Ltd.

Paper IV

A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography

R. Martins, C. J. Maia, J. A. Queiroz, F. Sousa

Journal of Separation Science. 2012. 35(22): 3217-3226

Rita Martins
Cláudio J. Maia
João A. Queiroz
Fani Sousa

CICS-UBI—Health Sciences Research Centre, University of Beira Interior, Covilhã, Portugal

Received March 30, 2012
Revised May 28, 2012
Accepted June 02, 2012

Research Article

A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography

The relevance of RNA in many biological functions has been recognized, broadening the scope of RNA research activities, from basic to applied sciences, also aiming the translation to clinical fields. The preparation and purification of RNA is a critical step for further application, since the quality of the template is crucial to ensure reproducibility and biological relevance. Therefore, the establishment of new tools that allows the isolation of pure RNA with high quality is of particular importance. New chromatographic strategies for RNA purification were considered, exploiting affinity interactions between amino acids and nucleic acids. In the present study, a single arginine-affinity chromatography step was employed for the purification of RNA from a total eukaryotic nucleic acid extract, thus eliminating several steps compared with current RNA isolation procedures. The application of this process resulted in a high RNA recovery yield of 96 ± 17% and the quality control analysis revealed a high integrity (28S:18S ratio = 1.96) in RNA preparations as well as a good purity, demonstrated by the scarce detection of proteins and the reduction on genomic DNA contamination to residual concentrations. Furthermore, the performance of the new RNA isolation method was tested regarding the applicability of the isolated RNA in modern molecular biology techniques. Hence, this new affinity approach will simplify the isolation and purification of RNA, which can bring great improvements in biomedical investigation.

Keywords: Affinity chromatography / Arginine / Gene expression / Real-time PCR / RNA isolation
DOI 10.1002/jssc.201200338

1 Introduction

The increasing awareness that RNA is one of the central molecules in cellular processes is broadening the RNA-based research [1]. Understanding the role of the various types of RNA present in cellular events is critical to develop new methods of diagnosis and treating diseases [2]. On the other hand, RNA is also the starting material in numerous molecular biology procedures involving the characterization of known messenger RNA (mRNA), the identification of unknown genes, and the assignment of function to several proteins. Moreover, in applications such as complementary DNA (cDNA) library construction, Northern blot analysis, reverse transcription (RT), and *in situ* hybridization analysis, a major factor determining the rate of success is the quality of initial RNA. Although RNA purification is a first critical step of a number of preparative and analytical methods, the commonly used isolation techniques present several limitations and have changed little in the past years [3]. To overcome this issue, the development of new tools that allows the isolation of biological and chemically stable RNA is of particular importance. The mainly used procedure for total RNA isolation employs a combination of denaturing agents, acid phenol chloroform extraction followed by precipitation of the nucleic acids [3, 4]. This procedure has the disadvantages of using hazardous products and of being very time consuming and highly operator dependent.

Liquid chromatography, especially HPLC, has been largely applied in attempting to overcome the limitations on RNA purification [5–9]. In these cases, the RNA is prepared through linear plasmid DNA (pDNA) templates for large-scale *in vitro* transcription and the HPLC methods, mainly reversed-phase, size-exclusion, or anion-exchange chromatography, are used to separate the desired RNA oligonucleotide from the transcription mixture. More recently, a robust affinity-purification protocol via ribozyme-cleavable RNA affinity purification tags was also proposed

Correspondence: Dr. Sousa, CICS-UBI—Health Sciences Research Centre, University of Beira Interior, Avenida Infante D. Henrique, 6200-506 Covilhã, Portugal
E-mail: fani.sousa@fcsaude.ubi.pt
Fax: +351 275 329 099

Abbreviations: AC, affinity chromatography; **BCA**, bicinchoninic acid; **cDNA**, complementary DNA; **Ct**, threshold cycle; **DEPC**, diethylpyrocarbonate; **gDNA**, genomic DNA; **hGAPDH**, human glyceraldehyde 3-phosphate dehydrogenase; **ncRNA**, noncoding RNA; **PC-3 cells**, human prostate cancer cell line; **pDNA**, plasmid DNA; **qPCR**, real-time PCR; **RT**, reverse transcription; **rRNA**, ribosomal RNA; **RNases**, RNA-degrading enzymes; **STD-NMR**, saturation transfer difference-nuclear magnetic resonance; **SPR**, surface plasmon resonance; **sc**, supercoiled; **tRNA**, transfer RNA

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[10]. In those applications, *in vitro* RNA transcripts are used as substitutes for native RNAs. However, most of the native RNAs have posttranscriptional modifications [11] and some of the modifications are quite important for their structure and function. Thus, isolation of intact RNAs from cells is essential for their study and application.

In this work, affinity chromatography (AC) with immobilized arginine is considered a potential technique for RNA isolation from a biological source, because of its unique characteristic of using a biospecific ligand to purify biomolecules, on the basis of their biological function or individual chemical structure. In fact, the use of amino acids as immobilized ligands for AC has been recently exploited and implemented, by our research group, as an effective methodology for nucleic acids purification [12–14]. Histidine [15,16] and arginine [17–19] have been used as amino acid ligands, and their ability to isolate supercoiled (sc) pDNA proved the presence of specific interactions occurring between pDNA and the amino acid based matrices. Furthermore, histidine AC was also applied in the purification of RNA [20,21]. As a result, histidine matrix showed a specific recognition for 6S RNA, allowing its purification from a complex mixture of *Escherichia coli* regulatory noncoding RNA (ncRNA) molecules [20]. Subsequently, the simultaneous isolation of ncRNA and ribosomal RNA (rRNA) was also accomplished using histidine AC. In this strategy, both RNA classes were accurately purified from *E. coli* impurities (genomic DNA (gDNA) and proteins) [21].

On the other hand, the application of arginine matrix showed some improvements in pDNA purification over other chromatographic techniques [13]. The arginine-based support allowed the efficient separation of plasmid isoforms, revealing the presence of a specific recognition for sc isoform [22]. Moreover, the different interactions of arginine ligands with pDNA, RNA, and gDNA suggested its potential application for the selective recovery of any nucleic acid. Additionally, it was found that the simplified purification process achieved with this support had a significant impact on sc pDNA stability, enhancing its biological function [18].

Building on the interesting results obtained, we will explore the possibility of using arginine AC to selectively isolate total RNA with high quality, in view of the application in molecular biology procedures, namely for gene expression analysis. This work intends to aid in the development of new procedures for RNA isolation and purification, which are generally recognized to be crucial for the overall success of RNA-based analyses [23,24]. Therefore, it is expected that this new affinity protocol for RNA isolation can offer advantages over other less-selective and time-consuming multistep procedures and can improve process economics.

2 Materials and methods

2.1 Materials

Human caucasian prostate adenocarcinoma cell line (PC-3, ECACC 90112714) was purchased from the European Collection of Cell Cultures (ECACC, Salisbury, UK). Cell culture reagents, namely RPMI 1640, and trypsin/EDTA were purchase from Sigma (St. Louis, MO, USA). Fetal bovine serum was obtained in Biochrom (Berlin, Germany) while penicillin/streptomycin solution was obtained from Invitrogen (Carlsbad, CA, USA). All the chemicals used in the cell lysis buffer were obtained from Sigma. Arginine–Sepharose 4B gel was obtained from GE Healthcare Biosciences (Uppsala, Sweden). The compounds used for chromatographic experiments were sodium chloride purchased from Panreac (Barcelona, Spain) and Tris base from Fisher Scientific (Leicestershine, UK). Glycogen was obtained from USB (Cleveland, OH, USA). All solutions were freshly prepared using 0.05% diethylpyrocarbonate (DEPC)-treated water from Fluka (Sigma) and the elution buffers were filtered through a 0.20-μm pore size membrane (Schleicher Schuell, Dassel, Germany) and degassed ultrasonically. The DNA molecular weight marker, HyperLadder I, was obtained from Bioline (London, UK). All the experiments were performed with RNase-free disposables.

2.2 Methods

2.2.1 Cell culture and lysis

PC-3 cells were initially cultured in 25-cm^2 flasks in RPMI 1640 at 37°C in a humidified incubator in 95% air/ 5% CO_2. The culture medium was supplemented with 10% (v/v) fetal bovine serum and 1% (v/v) penicillin/streptomycin. The PC-3 cells were routinely passaged at 90–95% confluence to 75-cm^2 flasks. In order to obtain total nucleic acid extracts, PC-3 cells from passages number 19, 21, and 22 were collected by washing and detaching with 0.25% trypsin/EDTA. The trypsin solution was neutralized with RPMI 1640 supplemented culture medium and the cells were spun down with a centrifuge (Sigma 3K18C, Bioblock Scientific; Sigma Laboratory Centrifuges) at 1500 rpm for 5 min at room temperature, and resuspended in culture medium. The cells spin and resuspension was repeated to ensure complete removal of trypsin. The recovered cells were resuspended in 1 mL of PBS and the total cells were counted using a Neubauer chamber. Next, approximately 10×10^6 cells were spun down and the pellets were resuspended in 2 mL denaturing cell lysis solution (4 M guanidinium thiocyanate; 25 mM sodium citrate, pH 4.0; 0.5% (m/v) *N*-laurosylsarcosine; and 0.1 M β-mercaptoethanol). After 5-min incubation at room temperature, 2 mL of ice-cold isopropanol was added. The lysate solution was homogenized by inversion and incubated on ice for 5 min. The precipitated molecules were recovered by centrifugation at 16 000 × *g* for 20 min at 4°C. The pellet was washed with 1 mL of 75% ethanol in DEPC-treated water and incubated at room temperature for 10 min, followed by a 5-min centrifugation at 16 000 × *g* (4°C). The air-dried pellet was dissolved in 240 μL of 0.05% DEPC-treated water and incubated for 5 min at room temperature to ensure

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

complete solubilization. The concentration of total nucleic acid preparation was estimated using NanoPhotometer (IMPLEN, Munich, Germany).

Besides the preparation of total nucleic acid extracts, total RNA was also isolated by using a conventional procedure. Thus, approximately 10^7 PC-3 cells from the same cell passages mentioned above were treated with commercial TRI reagent (Ambion, Carlsbad, CA, USA), according to manufacturer's instructions.

2.2.2 Preparative arginine–agarose chromatography

Chromatographic experiments were carried out using an ÄKTA Avant system with UNICORN 6 software (GE Healthcare). It was used a commercial arginine–Sepharose 4B gel characterized by the manufacturer as a cross-linked 4% beaded agarose matrix with a 12-atom spacer and an extent of labeling between 14 and 20 μmoL/mL. The stationary phase was packed in a 10 mm diameter × 20 mm long (∼2 mL) column. Temperature was maintained at 7°C during the experiments by a circulating water bath.

The stationary phase was equilibrated with 190 mM NaCl in 10 mM Tris-HCl (pH 8.0) buffer at a flow rate of 1 mL/min. Conductivity was controlled at 19 mS/cm. With relation to the sample application, 200 ng/μL of total nucleic acid preparation, from passages 19, 21, and 22, was injected onto the column using a 100-μL loop at the same flow rate. After washing out the unbound material with 190 mM NaCl in 10 mM Tris-HCl (pH 8.0) buffer, the ionic strength of the buffer was increased to 1 M NaCl in 10 mM Tris-HCl (pH 8.0) buffer. The different peaks were monitored with a UV detector at 260 nm. Each peak was automatically collected in a climate-controlled fraction collector chamber and was concentrated by addition of 20 μg glycogen and four volumes of 100% ethanol. After a 2-h incubating period at –80°C, the fractions were recovered by centrifugation at 16 000 × *g* for 20 min at 4°C. Pellets were air-dried for 15 min and reconstituted in DEPC-treated water. After chromatographic runs, the column was cleaned with three column volumes of 0.2 M NaOH.

For the identification of eluting species and evaluation of RNA integrity, the samples were resolved on a 1.2% native agarose gel in TAE buffer (40 mM Tris base, 20 mM acetic acid, and 1 mM EDTA, pH 8.0) prepared in DEPC-treated water. The gels were stained with ethidium bromide (0.5 μg/mL), and photographed. The assessment of RNA quality and downstream analysis were further performed as described below.

2.2.3 RNA quantification

ÄKTA Avant system (GE Healthcare) with arginine–agarose column was also used to quantify the RNA present in each sample recovered from the chromatographic purification. A calibration curve was prepared with RNA standards (10–80 μg/mL) purified with TRI reagent (Ambion). The standard experiments were undertaken in triplicate. The experiments were performed by injecting 100 μL of RNA standards onto arginine–agarose matrix after an equilibration step with 190 mM NaCl in 10 mM Tris-HCl (pH 8.0) buffer, using a flow rate of 1 mL/min. The elution buffer of 1 M NaCl in 10 mM Tris-HCl (pH 8.0) buffer was immediately applied to the column in order to favor entire RNA elution. The areas of the peaks obtained in the ÄKTA histogram were calculated using UNICORN 6 software. A standard curve was obtained by a linear fit between the sample concentration and peak integration area, with a correlation coefficient of 0.998. The RNA quantification in other samples was accomplished by comparing their peak areas with the respective standard curve.

2.2.4 Protein analysis

Protein residual contamination in RNA samples, either collected from the purification with arginine–agarose support or isolated by TRI reagent, was assessed by using the micro-BCA (bicinchoninic acid) assay (Thermo Fisher Scientific Inc., Rockford, IL, USA), according to manufacturer's instructions. Briefly, the calibration curve was prepared using BSA standards (0.01–0.1 mg/mL). A total of 50 μL of each standard or RNA samples was added to 200 μL of BCA reagent in a microplate and incubated for 30 min at 60°C. Absorbance was measured at 570 nm in a microplate reader.

2.2.5 RT-PCR and real-time PCR (qPCR)

All RNA samples in study were amplified by RT-PCR in a thermo cycler (Biometra, Goettingen, Germany). cDNAs synthesis was performed using RevertAid First Strand cDNA Synthesis Kit (Fermentas, Thermo Fisher Scientific Inc.), according to manufacturer's instructions. A total of 1 μg of total RNA was used to initiate cDNA synthesis. PCR reactions were carried out using 1 μL of the synthesized cDNAs in a 25 μL reaction containing 1 × Taq DNA polymerase buffer (Xnzytech, Lisboa, Portugal), 500 μM deoxynucleotide triphosphates (Amersham Biosciences, Uppsala, Sweden), 3 mM of magnesium chloride (Fermentas, Thermo Fisher Inc.), 300 nM of each primer, and 0.125 U of Taq DNA polymerase (Xnzytech). The used human glyceraldehyde 3-phosphate dehydrogenase (hGAPDH) primers were provided by RevertAid First Strand cDNA Synthesis Kit, which amplify products of 496 bp. The cycling conditions were used in accordance to the instructions of cDNA synthesis kit in Control PCR amplification section. To confirm the presence and purity of amplicons, RT-PCR products were analyzed by 1% agarose gel electrophoresis.

gDNA contamination and gene expression was evaluated using total RNA purified with the arginine matrix or TRI reagent in an iQ5 Multicolor Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA). Specific primers to hGAPDH and 18S rRNA were used in gDNA and cDNAs amplification according to Table 1. The qPCR efficiency was determined for all primer sets using serial dilutions of cDNA samples (1:1, 1:10, 1:100, and 1:1000). A standard curve was generated by serial dilution of PC-3 gDNA (purified with the Wizard gDNA purification kit; Promega,

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 1. Technical features of the specific primers used in qPCR analysis

Primer	Sequence		Amplified fragment (bp)	Annealing temperature (°C)
	Sense	Antisense		
hGAPDH_intron[a)]	5′-CCCACACACATGCACTTACC-3′	5′-CCCACCCCTTCTCTAAGTCC-3′	176	60
hGAPDH_74 hGAPDH_149	5′-CGCCCGCAGCCGACACATC-3′	5′-CGCCCAATACGACCAAATCCG-3′	75	
18S _980 18S _1128	5′-AAGACGGACCAGAGCGAAAG-3′	5′-GGCGGGTCATGGGAATA-3′	148	

a) The hGAPDH primers inside the first gene intron were used for gDNA amplification.

Madison, WI, USA) in the range of 0.8–800 ng/μL. qPCR reactions were carried out using 1 μL of gDNA standards or synthesized cDNA in a 20 μL reaction containing 10 μL Maxima SYBR Green/ROX qPCR Master mix (Fermentas, Thermo Fisher Scientific Inc.) and 300 nM hGAPDH or 18S primers. All reactions were performed in 96-well plates for PCR heat-sealed with heat sealing film (Bio-Rad Laboratories). After an initial denaturation at 95°C for 5 min, cycling conditions were as follows: 35 cycles consisting of denaturation at 95°C for 10 s, annealing at 60°C for 30 s, and extension at 72°C for 10 s. The amplified PCR fragments were checked by melting curves: reactions were heated from 55 to 95°C with 10 s holds at each temperature (0.05°C/s). Samples were run in triplicate for each assay. Results were analyzed using iQ5 optical system software version 2.0 after manual adjustment of the baseline and fluorescence threshold. Fold differences were calculated following the mathematical model proposed by Pfaffl using the formula $2^{-\Delta Ct}$ [25].

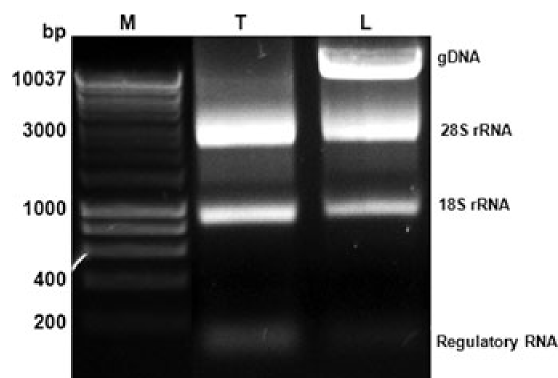

Figure 1. Agarose gel electrophoresis of the nucleic acids in PC-3 total nucleic acid extract recovered by guanidinium-based lysis (lane L) and in samples isolated with TRI reagent (lane T). Lane M, DNA molecular weight marker. Data are representative of three independent experiments with three cell passages.

3 Results and discussion

3.1 RNA purification from PC-3 total nucleic acid extracts

In this work, the applicability of a method based on arginine AC for the isolation of RNA from PC-3 cells, a well-established prostate cancer cell line [26], was evaluated.

Total nucleic acid preparations were obtained by chemical lysis with guanidinium buffer, which due to its chaotropic effect disrupts the plasma membrane and induces organelle lysis. This process liberates heterogeneous nuclear RNA and gDNA from nucleus and mitochondria, all of which are recovered with the cytoplasmic RNA. In addition, guanidinium buffer causes efficient unfolding of proteins, by which RNA-degrading enzymes (RNases) tertiary structure is distorted, inhibiting their activity [27]. Therefore, it is not necessary to add additional RNase inhibitors.

Figure 1 shows the electrophoretic profile of the nucleic acids present in a PC-3 total nucleic acid preparation after chemical lysis (Fig. 1, lane L). This sample contains four distinct nucleic acid species, which correspond to gDNA, 28S rRNA, 18S rRNA, and low molecular weight RNA species, from the slower to faster migrating biomolecules. Total RNA isolated by TRI reagent was also loaded onto the gel as a control (Fig. 1, lane T). The electrophoretic analysis indicates an accurate banding profile for RNA molecules, with a clear definition of 28S and 18S rRNA species. The greater fluorescence of the 28S rRNA demonstrates the integrity of the samples. All transfer RNA (tRNA) and the low molecular weight 5S and 5.8S rRNA species comigrate and appear at the bottom of the gel. Because electrophoretic analysis was performed under native conditions, a resultant smear can be occasionally visualized. The smear is not due to sample degradation but instead it reflects persistent RNA secondary structure [27]. Therefore, this qualitative analysis suggests that RNA in total nucleic acid extracts is chemically intact.

Total RNA isolation with arginine-chromatography was achieved after several optimizing experiments using different salt concentrations, buffer types, and temperatures in order to select the best conditions for RNA binding and elution (results not shown). Therefore, mild chromatographic conditions using low NaCl concentrations in Tris buffer pH 8.0 were used while maintaining the temperature at 7°C to prevent RNA degradation. Total RNA retention was achieved with NaCl concentrations between 150 and 200 mM and its elution occurred when using 1 M NaCl. Curiously, when no salt was present in the Tris buffer, gDNA and rRNA were

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

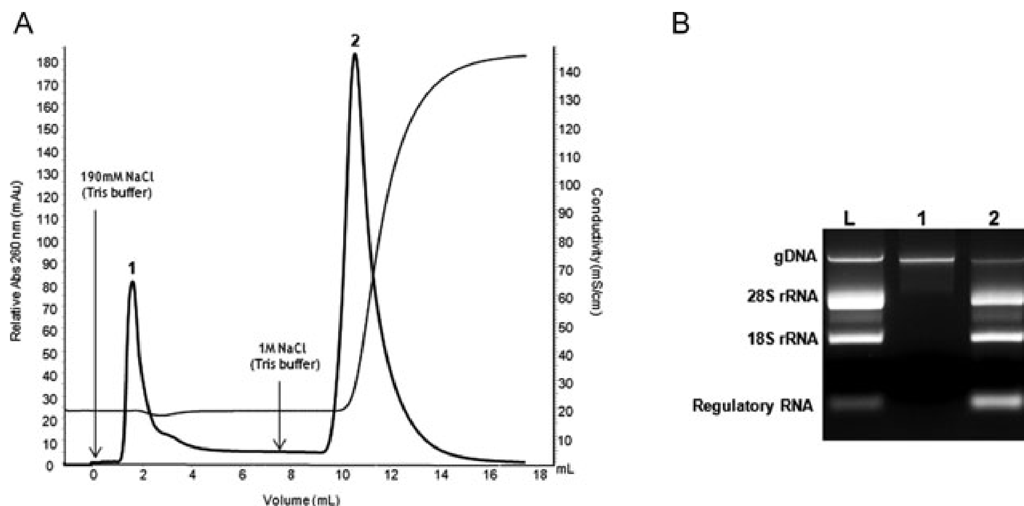

Figure 2. (A) Chromatographic profile of the purification of total RNA from PC-3 total nucleic acid extract by arginine–agarose chromatography. Elution was performed at 1.0 mL/min by stepwise increasing NaCl concentration in the eluent from 190 mM to 1 M, as represented by the arrows. The conductivity was followed along the chromatographic purification as indicated by the dashed line. (B) Agarose gel electrophoresis of the samples collected at the column outlet. Fractions corresponding to peaks (1) and (2) are shown in lanes 1 and 2, respectively. Lane L, total nucleic acid extract injected onto the column. Data are representative of three independent experiments with three cell passages.

observed to promptly elute, while sRNAs were retained. During these experiments, the need for strict control of chromatographic conditions was verified in order to maintain the reproducibility since a slight variation in conductivity (salt concentration and/or temperature) affected total RNA retention.

Figure 2A shows the chromatographic profile obtained after the injection of total nucleic acid preparation on the arginine support. The chromatographic run was initiated at low ionic strength with 190 mM NaCl in 10 mM Tris buffer (pH 8.0). Under these conditions, it was observed a flow-through peak containing unbound species. The elution of highly bound species was then achieved by increasing the ionic strength of the buffer to 1 M NaCl. The presence of two different peaks in the chromatogram is a consequence of the different interaction that nucleic acids exhibit with arginine–agarose matrix. The agarose gel electrophoresis was used to identify the different nucleic acid species from each peak (Fig. 2B). The electrophoretic profile presented in lanes 1 and 2 corresponds to the samples pooled from the respective peaks in the chromatogram. The total nucleic acid preparation injected on arginine matrix (Fig. 2B, lane L) was also run in the gel for comparative purposes. Hence, electrophoretic analysis showed that the first peak of unbound species corresponds to gDNA (lane 1), while the second peak mainly refers to RNA species (lane 2). These results suggest that the different functional classes of RNA present a stronger interaction with the arginine matrix than gDNA.

Although the second peak mainly includes RNA species, a slight band of gDNA is still visible in the electrophoretic analysis (lane 2). Thus, it is important to determine the extent of gDNA contamination on RNA samples. Further experiments on the quantification of gDNA by qPCR were performed and they are discussed in the next section of RNA quality characterization.

Additionally, it was attempted to improve the quality of RNA preparations by performing some changes in the elution condition on arginine AC.

In AC, the elution of a target solute that is bound to the affinity ligands can be achieved through addition of a competing agent in the elution buffer rather than changing the ionic strength, pH, or polarity of elution buffer. In this study, the competitive studies were performed by adding 250 mM of arginine to the elution buffer in linear or stepwise gradients (data not shown). Arginine was used as competing agent to exploit specific elution of RNA from the column and therefore to evaluate the possibility to reach higher purification factors. However, the experiments did not result in an improvement in RNA purification, because higher quantity of gDNA was recovered in RNA fractions. These results suggest that the presence of arginine in the elution buffer has also an effect in gDNA elution (data not shown).

In AC, the interactions occurring between a target biomolecule and its specific ligand are responsible for the high selectivity achieved in this technique. Those interactions are similar to the contacts described in many biological systems [28]. Thus, the binding mechanism is suggested to involve phenomenological interactions, such as biorecognition, between the amino acid and RNA, including, electrostatic, hydrophobic interactions, multiple hydrogen bonds, dipole–dipole forces, or cation–π interactions [29]. However, depending on the environmental conditions established, some interactions can be more favored than others, becoming more evident under those conditions. Since RNA is negatively charged due to the phosphate groups in its backbone, it is reasonable to predict a favored electrostatic interaction between RNA

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

phosphate groups and arginine ligands. In fact, in some molecular recognition studies, arginine is reported as a preferential amino acid to contact with RNA when the overall negative charge of RNA is considered [29, 30]. Additionally, saturation transfer difference-nuclear magnetic resonance (STD-NMR) spectroscopy and surface plasmon resonance (SPR) biosensor techniques recently reported that adenine, cytosine, and guanine polynucleotides bind to arginine–agarose support mainly through the sugar-phosphate backbone [31].

On the other hand, gDNA was observed to be the less-retained nucleic acid, revealing a not so strong affinity for the arginine–agarose support as in case of RNA species. These findings are in agreement with some previous studies focused on the purification of pDNA by arginine chromatography [18, 32]. According to the authors, the negative charge of the biomolecules is important for their interaction with arginine, nonetheless the nucleotide bases exposure is also suggested to have a crucial role in nucleic acid retention [18, 32]. Thus, despite the negative charge of gDNA, its double-stranded structure causes the coverage of the nucleotides bases, disabling gDNA interactions. In line with this, the bases exposure on RNA species should play an important role in the favorable interaction found with arginine–agarose.

Furthermore, structural diversity of RNAs was recently described to be of significant importance in protein–RNA interactions because RNA can exhibit different moieties according to its folding state [1, 33]. Bioinformatics predictions showed that amino acids complexes with mRNA and tRNA, exhibiting less-compact secondary structures, have a greater number of base-specific contacts and fewer backbone contacts, while the amino acids complexes with rRNA (more compact secondary structures) have less base-specific contacts [33].

In this study, the purification approach described enabled the separation of gDNA and RNA. Moreover, the arginine support has shown ability to interact with all RNA classes even with different conformational rearrangements. The multiposition interaction of arginine with RNA sites [34, 35] can explain this result. The multiplicity of interaction sites can occur because arginine has two different polar centers with which RNA can strongly associate: at α-carbon group and the side chain guanidinium [29]. Thus, it is reasonable to suppose that the retention of all functional classes of RNA in arginine–agarose matrix is due to arginine side chain, which can promote multicontact with RNA backbone or RNA bases, according to RNA folding. Overall, it is suggested that although electrostatic interactions could play an important role on RNA retention, the bases contacts are also involved and modulate some favored interaction and specificity found in arginine–agarose chromatography.

3.2 RNA quality characterization

Fundamental criteria for extraction and purification procedures of total RNA concerning molecular biology application, in particular qPCR, have been reported [36, 37]. Accordingly, the RNA preparation should be free of proteins and gDNA, especially if the target is an intronless gene, should be undegraded (28S:18S ratio should be roughly between 1.8 and 2.0), free of enzymatic inhibitors for RT and PCR reactions, free of any substance that complexes essential reaction cofactors, such as Mg^{2+} or Mn^{2+}, and free of nucleases [36, 37].

Table 2. Total RNA quantification in different PC-3 total nucleic acid preparations injected onto arginine column

Total nucleic acid preparation injected (μg)	Integrated peak area	Total RNA (μg)
30	247.91	23.35
35	257.49	24.35
40	273.15	25.97
50	355.02	34.49
155		108.20

To validate these criteria, RNA samples obtained by arginine affinity purification were quantified, the RNA integrity was evaluated by ribosomal band intensity, and the presence of gDNA and proteins was assessed. Total RNA control samples obtained by using the commercial TRI reagent were also included in the measurements.

The results of RNA quantification are shown in Table 2 that presents the concentration of total RNA in different amounts of PC-3 nucleic acid extracts, which were applied to arginine chromatography. These results indicated that arginine–agarose matrix allowed the quantitation of different RNA contents in the complex mixture. In addition, the mass reduction relative to total nucleic acid extract corresponds to gDNA that is separated from total RNA by arginine–agarose matrix. Total RNA quantification by peak integration provided a reliable and accurate method because spectrometric methods often fail in sensitivity and are highly variable leading to over- or underestimation of the real RNA concentration [37, 38]. Regarding the recovery yields obtained, it was verified that from the chemical lysis of around 10^7 PC-3 cells, it was possible to obtain about 282 μg of nucleic acids. In Table 2, it can also be seen that from 114 $\pm$ 19 μg of total RNA present in the nucleic acids extract, 108 $\pm$ 11 μg of pure RNA was obtained after arginine AC purification, achieving a recovery yield of 96 $\pm$ 17% (Table 3). This RNA recovery is similar to the one obtained with TRI reagent, which enables the recovery of an RNA averaging amount of 157 $\pm$ 17 μg from the same starting number of cells.

The RNA integrity was assessed by agarose gel electrophoresis using UVIband–1D gel analysis software (Uvitec, Cambridge, UK) to determine the intensity of ribosomal bands. The proportion of the ribosomal bands (28S:18S) is crucial to guarantee RNA integrity, followed by the absence of low molecular weight RNA-degradation products. Figure 2B lane 2 shows the electrophoretic analysis of the RNA sample obtained after arginine AC isolation and Fig. 3 shows the semiquantitative analysis of the intensity of ribosomal bands.

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 3. Quantitative analysis of purity and recovery yield of RNA isolated by arginine–agarose chromatography. The correlation coefficients of gDNA and protein calibration curves were 0.995 and 0.993, respectively. Data are presented as means with SD ($n = 3$). TNA, total nucleic acid extract; Arg, arginine

Method	Sample	Volume (μL)	Nucleic acids (μg)		Recovery yield (%)	Proteins (ng/μL)	
			gDNA	RNA		200 TNA	400 TNA
Arg–AC	TNA extract	120	155 ± 21			18 ± 6	42 ± 4
			41 ± 29	114 ± 19			
	Peak 2 (RNA)		1 ± 0.27	108 ± 11	96 ± 17	ND[b)]	5 ± 2
TRI[a)]	RNA	120	0.39 ± 0.44	157 ± 17	____	ND[b)]	4 ± 2

a) Control sample.
b) ND, not detected by micro-BCA assay.

The ratio of 1.96 indicates that the intensity of 28S rRNA band is two times higher than the intensity of 18S rRNA, demonstrating a good RNA integrity. In fact, the integrity of the RNA molecules is a key factor for the overall success of further application. Furthermore, RNA is one of the most difficult materials to separate under chromatographic conditions. One reason for this is that RNA is degraded very quickly in nature and its stabilization is very difficult. Hence, the final RNA integrity will depend on maintaining the stability of the sample before separation, throughout the purification process, and also during the recovery of RNA fractions when the separation has been completed. Our results demonstrate that the isolation steps involved in the affinity procedure allowed maintaining the RNA stability. As previously discussed, the chaotropicity of guanidinium buffers inhibits RNase action, preventing RNA degradation. Additionally, the application of arginine AC can also be strongly associated with the preserved integrity observed in RNA samples since arginine, owing to its multiplicity for interactions, has been largely associated with stabilizing effects on RNA conformations [35, 39, 40].

With relation to the residual contamination of gDNA, it is an inherent problem during RNA purification due to the similar physicochemical properties of RNA and DNA. Quantification by qPCR indicated that the residual concentration of gDNA was of 8 ± 0.27 and 3 ± 0.44 ng/μL in RNA samples after arginine AC and in total RNA samples isolated by TRI reagent, respectively (Table 3). Nevertheless, arginine AC allowed a significant decrease in gDNA levels in total RNA preparations. As mentioned above, the double-stranded structure of gDNA does not favor the interaction with the arginine matrix, so reduced gDNA content on total RNA fraction would be expected.

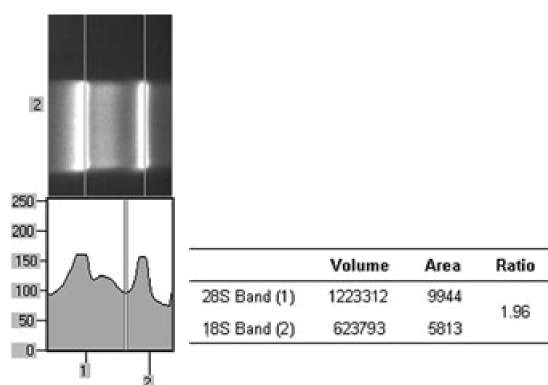

Figure 3. Semiquantitative analysis of the intensity of rRNA bands using UVIband–1D gel analysis software. 28S:18S ratio was calculated using the peak volumes. Data are representative of three independent experiments.

Some of the concerns associated with the presence of gDNA in RNA preparations are related to the interfering effect that it can promote on several molecular biology analyses, particularly in qPCR analysis. However, the impact of gDNA contamination on qPCR signals is often dependent on primer design strategy [41].

The protein analysis (Table 3) performed by micro-BCA method showed that when 200 ng/μL of total nucleic acid preparation was injected on arginine–agarose column, no protein was detected in total RNA or gDNA pools, indicating that protein content is inferior to the detection limit of the method (<5 ng/μL). However, a residual protein concentration of 5 ± 2 ng/μL was determined in total RNA samples, when the injection of total nucleic acid preparation was increased to 400 ng/μL. These results may indicate that proteins have a propensity to elute steadily throughout the gradient, which is probably due to the heterogeneous proteins content, promoting different interactions with arginine–agarose matrix. In line with these results, total RNA samples isolated with TRI reagent demonstrated undetectable proteins in concentrations of RNA up to 400 ng/μL, while above this concentration the residual protein level was of 4 ± 2 ng/μL.

Overall, regarding the described criteria for an accurate RNA extraction and purification method [36, 37], the results obtained so far demonstrate that the arginine affinity based protocol described here is a potential methodology for RNA isolation. The chemical lysis with guanidinium buffers used in this method is extensively described in literature for RNA extraction, including in many commercial kit and reagents, with no significant implications in downstream applications [27, 42]. Moreover, total nucleic acid extract is purified by AC controlled by an automatic system that improves reproducibility. In addition, arginine chromatography makes use

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

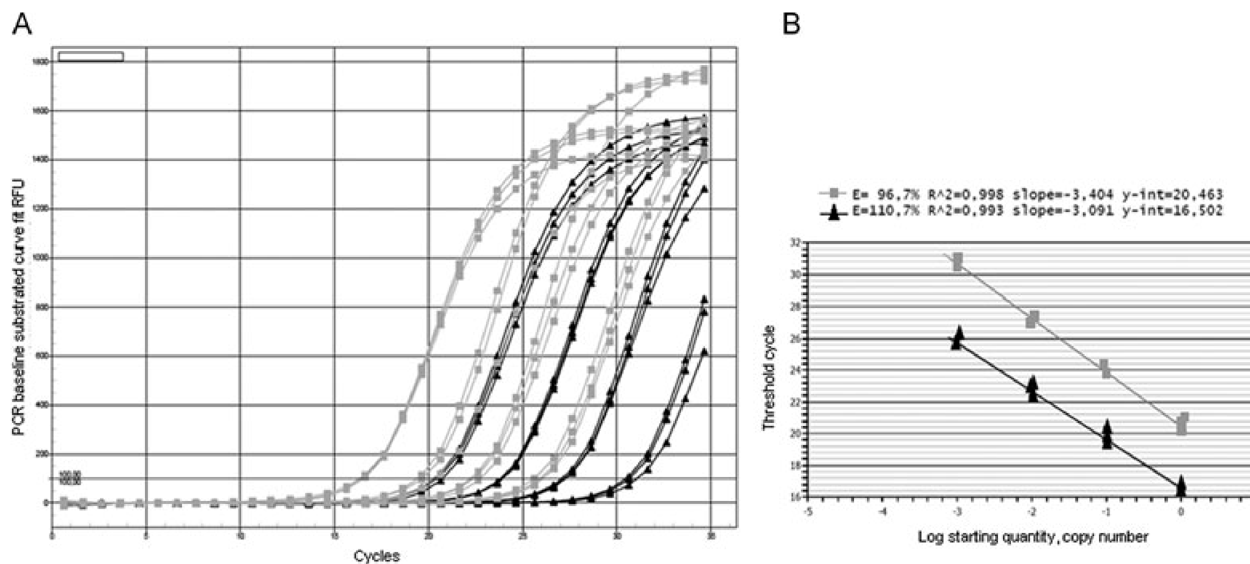

Figure 4. qPCR output showing amplification plots of hGAPDH expression (A) and PCR efficiencies (B). The squares and triangles lines represent hGAPDH expression profile from total RNA samples purified by arginine AC or TRI reagent, respectively. Data in (A) and (B) are representative of three separated experiments with total RNA samples isolated from three cell passages using both methods and were confirmed in other three experiments with 18S gene.

of low NaCl concentrations, rather than organic or toxic compounds, and the use of enzymes is not necessary. Finally, RNA concentration step employs ethanol precipitation in the presence of glycogen, which is used as a carrier agent to increase the recovery of RNA and to help in the visualization of the pellet. Ethanol is easily removed by centrifugation and evaporation and the use of 20 μg of glycogen is described not affecting downstream analysis [43]. Furthermore, total RNA samples were found with good integrity, low gDNA content, and the protein content was negligible.

Although chromatography is used more often in RNA isolation methodologies [6, 7, 10, 44], this is the first attempt to purify total RNA from eukaryotic cells using AC in an agarose support. Thus, it is considered that the implementation and optimization of this methodology can bring new insights to RNA purification.

3.3 Evaluation of total RNA isolated by arginine AC

With the aim to characterize arginine affinity based method for RNA isolation, total RNA samples obtained were used as template in qPCR, which is one of the most widely used techniques in modern molecular biology. Since TRI reagent is extensively used to isolate RNA for gene expression analysis, total RNA samples extracted by TRI reagent were used as a reference group.

First, the cDNA synthesis was successfully carried out using total RNA isolated by arginine column or TRI reagent. The quality of cDNAs was evaluated by conventional PCR, which enabled the amplification of hGAPDH fragments with 496 bp in both RNA samples (data not shown).

In order to evaluate if there is significant differences in detection of gene expression levels between RNA samples, the mRNA expression of hGAPDH and 18S genes was analyzed because they are two common housekeeping genes often used as endogenous references in qPCR [45, 46].

Figure 4 shows the qPCR plots (A) and PCR efficiency slops (B) for hGAPDH expression profiling from cDNAs synthesized from total RNA samples isolated by arginine affinity based method or TRI reagent. The amplification plots allowed obtaining the threshold cycle values (Ct), while PCR standard curves demonstrate the primers efficiencies. The spaced amplification curves from each sample produced a linear standard curve with reaction efficiencies between 90 and 110%, which is indicative of a good efficiency [47].

The Pfaffl method ($2^{-\Delta Ct}$) is one of the most recurrent approaches for calculating relative gene expression and it gives the possibility to present the expression data as "fold variation" [48, 49].

Table 4 indicates the mean of Ct values obtained for each sample and the fold variation between the two methods used

Table 4. Ct values obtained from the qPCR plots of each housekeeping gene and calculations of fold variation in expression between the two methods used for total RNA isolation. Data are presented as means with SD ($n = 3$)

Gene	Method	Ct	$2^{-\Delta Ct}$
GAPDH	Arg–AC	21 ± 3	16
	TRI	17 ± 2	
18S	Arg–AC	12 ± 4	4
	TRI	10 ± 3	

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

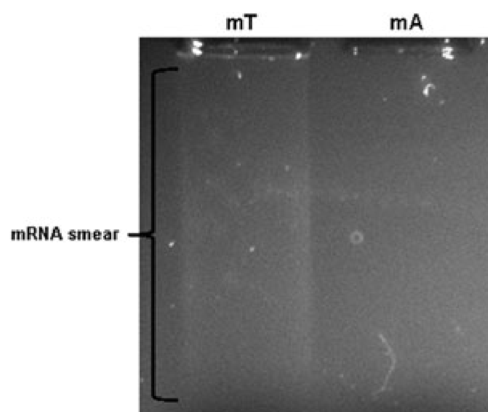

Figure 5. Electrophoretic analysis of mRNA samples isolated by Oligotex mRNA Mini Kit from total RNA purified by arginine AC and isolated by TRI reagent. Lane mT, mRNA from total RNA samples isolated by TRI; lane mA, mRNA from total RNA samples isolated by arginine–agarose column. Data were confirmed in three independent experiments with total RNA samples isolated from three cell passages using both methods.

to extract total RNA. The hGAPDH or 18S gene expression was found to be diminished in samples where total RNA was isolated by arginine affinity method, showing different fold variations between genes. The number of hGAPDH mRNA molecules in total RNA samples isolated by arginine AC is decreased by 16-fold relatively to hGAPDH mRNA molecules isolated by TRI reagent, while the expression of 18S rRNA is decreased by fourfold. These expression differences might be due to the fact that the methods may isolate different proportion of RNA molecules, leading to a higher concentration of rRNA comparatively to mRNA in total RNA samples.

Thus, it was supposed that Ct values obtained by qPCR were probably affected by the quantities of the starting material, since the same amounts of RNA should be used when comparing different samples using qPCR [37]. Therefore, mRNA was isolated from total RNA samples (Oligotex mRNA Mini Kit, Qiagen Hilden, Germany) obtained either by arginine column or TRI reagent, in order to guarantee that the same initial concentration of molecules was present when performing qPCR. The mRNA concentration and quality were measured using NanoPhotometer and through agarose gel electrophoresis (Fig. 5). Curiously, no smear was visualized in mRNA sample obtained from arginine affinity based method (Fig. 5, lane mA), contrarily to what was observed in the other sample (lane mT). Next, hGAPDH amplification was quantified by qPCR in both mRNA samples, which revealed a similar fold variation comparatively to the experiments using total RNA samples isolated by arginine AC or TRI reagent (data not shown). These findings sustain the hypothesis that the expression differences can be due to different amounts of starting molecules. This is supported by the fact that mRNA molecules from arginine AC purification were not visualized in agarose gel, suggesting that the real RNA concentration was not the one measured in the spectrophotometer. Considering the 16-fold variation between the two methods for hGAPDH expression, the mRNA injected in the agarose gel was indeed less than the expected and it was below the detection limit of electrophoresis technique. It is possible that the mRNA quantified spectrophotometrically has been overestimated due to the presence of contaminants in the sample. Regarding the gDNA contamination in total RNA samples, a higher gDNA level in total RNA obtained by arginine affinity method than in TRI samples was found (Table 3). The primers used to detect hGAPDH did not amplify gDNA because the primers are located in different exons. On the other hand, 18S primers can amplify contaminating gDNA. Our results suggest that the residual gDNA in total RNA samples did not affect the gene amplification because the Ct values obtained from arginine affinity protocol are higher than the Ct values determined in total RNA extracted by TRI reagent. This means that more PCR amplification cycles in samples from arginine AC method were necessary than in TRI samples, in order to detect the presence of RNA molecules. If qPCR was affected by gDNA contamination, Ct values in the samples from arginine AC method would be lower than in TRI samples.

Nevertheless, the overall fold variation (<16) for the housekeeping genes used are considered minor differences in RNA concentrations [24], which encourage the use of arginine affinity based method to prepare RNA samples to be applied in downstream analysis aiming gene expression studies. Yet, the use of total RNA isolated by the arginine AC should be taken cautiously in qPCR analysis of genes with low-expression levels. Moreover, the accuracy of the amplification plots and PCR efficiencies obtained in qPCR quantification attest the performance of arginine affinity based method to isolate total RNA with high quality, since degraded or impure RNA can limit the efficiency of the RT reaction and reduce yield [23].

4 Concluding remarks

In the present study, a new affinity approach for total RNA isolation from PC-3 total nucleic acid extracts using arginine AC is introduced. Amino acid based AC has shown a great potential for the purification of nucleic acids, because it combines the selectivity of a naturally occurring biological interaction with the simplicity of a single small molecule used as a ligand. The exploitation of specific affinity interactions between RNA and arginine allowed obtaining total RNA preparations with high integrity and purity, which were attested by qPCR analysis.

The results of control analysis and performance indicated that the chromatographic separation is a promising strategy for total RNA isolation. Nevertheless, some experimental setting should be reevaluated in order to render arginine–agarose chromatography viable for the purification of total RNA pursuing gene expression analysis.

This work was supported by FCT, the Portuguese Foundation for Science and Technology (PTDC/EBB-BIO/114320/2009).

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Rita Martins also acknowledges a fellowship (SFRH/BD/64100/2009) from FCT.

The authors have declared no conflict of interest.

5 References

[1] Doudna, J. A., *Nat. Struct. Biol.* 2000, *7* (Suppl), 954–956.

[2] Sharp, P. A., *Cell* 2009, *136*, 577–580.

[3] Tan, S. C., Yiap, B. C., *J. Biomed. Biotechnol.* 2009, *2009*, 1–10.

[4] Vomelova, I., Vanickova, Z., Sedo, A., *Folia Biol.* 2009, *55*, 243–251.

[5] Azarani, A., Hecker, K. H., *Nucleic Acids Res.* 2001, *29*, E7.

[6] Dickman, M. J., Hornby, D. P., *RNA* 2006, *12*, 691–696.

[7] Kim, I., McKenna, S. A., Viani Puglisi, E., Puglisi, J. D., *RNA* 2007, *13*, 289–294.

[8] McCarthy, S. M., Gilar, M., Gebler, J., *Anal. Biochem.* 2009, *390*, 181–188.

[9] Easton, L. E., Shibata, Y., Lukavsky, P. J., *RNA* 2010, *16*, 647–653.

[10] Batey, R. T., Kieft, J. S., *RNA* 2007, *13*, 1384–1389.

[11] Dunin-Horkawicz, S., Czerwoniec, A., Gajda, M. J., Feder, M., Grosjean, H., Bujnicki, J. M., *Nucleic Acids Res.* 2006, *34*, D145–149.

[12] Sousa, F., Prazeres, D. M., Queiroz, J. A., *Trends Biotechnol.* 2008, *26*, 518–525.

[13] Sousa, A., Sousa, F., Queiroz, J. A., *J. Sep. Sci.* 2010, *33*, 2610–2618.

[14] Sousa, A., Sousa, F., Queiroz, J. A., *J. Chromatogr. B* 2009, *877*, 3257–3260.

[15] Sousa, F., Freitas, S., Azzoni, A. R., Prazeres, D. M., Queiroz, J. A., *Biotechnol. Appl. Biochem.* 2006, *45*, 131–140.

[16] Sousa, A., Sousa, F., Prazeres, D. M., Queiroz, J. A., *Biomed. Chromatogr.* 2009, *23*, 745–753.

[17] Sousa, F., Queiroz, J. A., *J. Chromatogr. A* 2011, *1218*, 124–129.

[18] Sousa, F., Prazeres, D. M. F., Queiroz, J. A., *J. Gene Med.* 2009, *11*, 79–88.

[19] Sousa, A., Sousa, F., Queiroz, J. A., *J. Sep. Sci.* 2009, *32*, 1665–1672.

[20] Martins, R., Queiroz, J. A., Sousa, F., *J. Mol. Recognit.* 2010, *23*, 519–524.

[21] Martins, R., Queiroz, J. A., Sousa, F., *Biomed. Chromatogr.* 2012 *26*, 781–788.

[22] Sousa, F., Prazeres, D. M., Queiroz, J. A., *Biomed. Chromatogr.* 2009, *23*, 160–165.

[23] Fleige, S., Pfaffl, M. W., *Mol. Aspects Med.* 2006, *27*, 126–139.

[24] Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller,R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., Wittwer, C. T., *Clin. Chem.* 2009, *55*, 611–622.

[25] Pfaffl, M. W., *Nucleic Acids Res.* 2001, *29*, e45.

[26] Tai, S., Sun, Y., Squires, J. M., Zhang, H., Oh, W. K., Liang, C.-Z., Huang, J., *Prostate* 2011, *71*, 1668–1679.

[27] Farrell, R. E., *RNA Methodologies*, Academic Press, New York 2005.

[28] Hage, D. S., *Clin. Chem.* 1999, *45*, 593–615.

[29] Yarus, M., Widmann, J. J., Knight, R., *J. Mol. Evol.* 2009, *69*, 406–429.

[30] Treger, M., Westhof, E., *J. Mol. Recognit.* 2001, *14*, 199–214.

[31] Cruz, C., Cabrita, E. J., Queiroz, J. A., *Anal. Bioanal. Chem.* 2011, *401*, 983–993.

[32] Sousa, F., Matos, T., Prazeres, D. M., Queiroz, J. A., *Anal. Biochem.* 2008, *374*, 432–434.

[33] Ellis, J. J., Broom, M., Jones, S., *Proteins* 2007, *66*, 903–911.

[34] Janas, T., Widmann, J. J., Knight, R., Yarus, M., *RNA* 2010, *16*, 805–816.

[35] Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D., Frankel, A. D., *Science* 1991, *252*, 1167–1171.

[36] Pfaffl, M. W., in: Worsfold, P. J., Gallagher, P. K. (Eds.), *Nucleic Acids, Encyclopedia of Analytical Science*, Academic Press, Elsevier Ltd., Philadelphia, USA 2005, pp. 417–426.

[37] Bustin, S. A., Nolan, T., *J. Biomol. Tech.* 2004, *15*, 155–166.

[38] Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A., Auffray, C., *Nucleic Acids Res.* 2005, *33*, e56.

[39] Jiang, F., Gorin, A., Hu, W., Majumdar, A., Baskerville, S., Xu, W., Ellington, A., Patel, D. J., *Structure* 1999, *7*, 1461–S1412.

[40] Puglisi, J. D., Wyatt, J. R., Tinoco Jr, I., *J. Mol. Biol.* 1990, *214*, 437–453.

[41] Laurell, H., Iacovoni, J. S., Abot, A., Svec, D., Maoret, J. J., Arnal, J. F., Kubista, M., *Nucleic Acids Res.* 2012, *40 (7)*, e51.

[42] Chomczynski, P., Sacchi, N., *Nature Protocols* 2006, *1*, 581–585.

[43] Hengen, P. N., *Trends Biochem. Sci.* 1996, *21*, 224–225.

[44] Perica, M. Ć., Šola, I., Urbas, L., Smrekar, F., Krajačić, M., *J. Chromatogr. A* 2009, *1216*, 2712–2716.

[45] Eisenberg, E., Levanon, E. Y., *Trends Genet.* 2003, *19*, 362–365.

[46] Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A., Heinen, E., *J. Biotechnol.* 1999, *75*, 291–295.

[47] Taylor, S., Wakem, M., Dijkman, G., Alsarraj, M., Nguyen, M., *Methods* 2010, *50*, S1–S5.

[48] VanGuilder, H. D., Vrana, K. E., Freeman, W. M., *Biotechniques* 2008, *44*, 619–626.

[49] Schmittgen, T. D., Livak, K. J., *Nat. Protocols* 2008, *3*, 1101–1108.

© 2012 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Paper V

New approach in RNA quantification using arginine-affinity chromatography: potential application in eukaryotic and chemically synthesized RNA

R. Martins, J. A. Queiroz, F. Sousa

Analytical and Bioanalytical Chemistry. 2013. 405(27): 8849-8858.

Anal Bioanal Chem (2013) 405:8849–8858
DOI 10.1007/s00216-013-7334-y

RESEARCH PAPER

New approach in RNA quantification using arginine-affinity chromatography: potential application in eukaryotic and chemically synthesized RNA

R. Martins · J. A. Queiroz · F. Sousa

Received: 11 June 2013 / Revised: 20 August 2013 / Accepted: 30 August 2013 / Published online: 14 September 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract The knowledge of RNA's role in biological systems and the recent recognition of its potential use as a reliable biotherapeutic tool increase the demand for development and validation of analytical methods for accurate analysis of RNA. Affinity chromatography is a unique technique because of the versatility of applications reliant on the affinity ligand used. Recently, an arginine-based matrix has been effectively applied in the purification of RNA because of the specific recognition mechanism for RNA molecules. This interaction is suggested to be due to the length of arginine side chain and its ability to produce good hydrogen bonding geometries, which promote multi-contact with RNA backbone or RNA bases, based on RNA folding. Thus, this work presents the development and validation of an analytical method with ultraviolet detection for the quantification of RNA using affinity chromatography with arginine amino acid as immobilized ligand. The method was validated according to International and European legislation for bioanalytical methods. The results revealed that the proposed method is suitable for the reliable detection, separation, and quantification of RNA, showing that the method is precise and accurate for concentrations up to 200 ng/μL of RNA. Furthermore, the versatility of the methodology was demonstrated by its applicability in the quantification of RNA from different eukaryotic cells and in crude samples of chemically synthesized RNA. Therefore, the proposed method demonstrates a potential multipurpose applicability in molecular biology RNA-based analysis and RNA therapeutics.

Keywords Affinity · Arginine · Chromatography · RNA · Transcription

R. Martins · J. A. Queiroz · F. Sousa (✉)
CICS-UBI – Health Sciences Research Centre, University of Beira Interior, Av. Infante D. Henrique, 6200-506 Covilhã, Portugal
e-mail: fani.sousa@fcsaude.ubi.pt

Introduction

RNA was recently recognized as a central molecule in cellular processes with implications in many diseases as well as in understanding of evolution, becoming one of the most exciting research areas of molecular biology [1]. From basic to applied research, many procedures employ pure and intact RNA molecules [2]. Working with low-quality and quantity RNA may strongly compromise the success of several procedures that are often labour-intensive, time-consuming, and highly expensive [3, 4]. Thus, the improvement of RNA isolation and purification methods, as well as the implementation of high-throughput analytical tools that can readily separate, quantify, and analyze RNA samples assumes an increasing significance, especially if diagnostic, therapeutic, or prognostic conclusions will be drawn.

Current methods for RNA quality assessment include spectrophotometric, gel electrophoresis, and other fluorescence-based quantifications [5, 6]. Ultraviolet (UV) absorption using a spectrophotometer is one of the most widely used methods for nucleic acid quantification because of its wide detection range, small sample volume, fast analysis, and low cost [7]. However, this method lacks specificity and sensitivity for RNA molecules. All nucleic acids (dsDNA, RNA, and ssDNA) absorb at 260 nm, and the method is not capable of distinguishing between the various forms of nucleic acid. Moreover, contaminants that absorb around 260 nm will also contribute to the final quantification, resulting in an overestimation of nucleic acid concentration. Agarose or acrylamide gel electrophoresis is another popular method for RNA analysis [8]. RNA concentration can be qualitatively measured by comparing the relative fluorescence intensity of unknown RNA bands to that of known RNA standards or quantitatively by using sophisticated equipment that use software to analyze a gel image, known as gel densitometry. Additionally, the contamination by other nucleic acids, such as genomic DNA (gDNA), can be assessed

as they can be visualized in distinct bands in the gel because of the different molecular weights [8]. Although the cost of using gel electrophoresis is relatively low, analysis requires significant amount of handling and additional precautions considering that fluorescent dyes are potentially hazardous. Thus, a proper handling and disposal of the reagents are required, especially with ethidium bromide, as it is a potential carcinogen [9]. Improving upon this approach, other intercalating fluorescent dye based systems, such as QuantiFluor™ RNA Dye System and Quant-iT™ RiboGreen® RNA Reagent, have been developed becoming a standard practice because of their higher sensitivity and more specific detection [10]. Nevertheless, inaccurate quantifications may occur as fluorescence dyes are not always specific for RNA and will bind to DNA as well. Recently, new instruments conjugating fluorescence and reverse transcriptase polymerase chain reaction (RT-qPCR) [11] and micro-fluidic capillary electrophoresis (2100 Bioanalyzer (Agilent Technologies), Experion (BioRad)) [12] were developed to minimize the limitations still existing in quality assessment of RNA samples. While these methods have unique and powerful features, the main disadvantage is cost, because specific instrumentation, reagents, and chips are required, which can make them inaccessible for most laboratories.

Chromatography is undoubtedly one of the most diverse and potent methods in biotechnology, both at analytical and preparative level. Ion pair reverse-phase chromatography has provided great advances in RNA characterization and analysis [13–17], nonetheless the environmental impact and cost that the use of organic compounds entails can be highly inconvenient for a biotechnological industry. Although chromatography is a reliable method with inherent simplicity and ease of operation, no other chromatographic-based techniques have been employed in RNA assessment so far.

Affinity chromatography is recognized as a powerful technique with great applicability in the purification of many biomolecules [18] because it exploits the principle of biomolecular recognition [19]. Recently, affinity chromatography was introduced in the development of new RNA purification processes [20, 21]. Newly, we have reported successful RNA purification approaches using amino acids as affinity ligands [22–24]. These studies were based on the fact that many different interactions exist between proteins and nucleic acids in biological systems, involving in particular basic amino acids such as histidine or arginine [25]. Moreover, several atomic and molecular recognition studies performed on RNA–protein interfaces have predicted preferential interactions between amino acids and nucleotide bases [26–28]. In fact, we have described the use of arginine amino acid as an affinity ligand for the development of a RNA purification procedure that enabled the isolation of pure and intact RNA with appropriate quality, proven by its use in gene expression analysis [24].

Given the simplicity, robustness, versatility, and high reproducibility of arginine affinity chromatography, the validation of an analytical method based on this technique, was considered imperative.

In light of the above, we developed and validated a simple, fast, and reliable analytical affinity chromatography method for the separation, identification, and quantification of total RNA. Furthermore, the potential applicability of the proposed approach was evaluated in cell lysates or total RNA samples obtained from different eukaryotic cell types and in crude samples of chemically synthesized RNA by in vitro transcription. The versatile application of the arginine affinity method can be significant in several standard molecular biology methods and in monitoring RNA bioproduct in therapeutic approaches.

Experimental section

Materials

Human Caucasian prostate adenocarcinoma cell line (PC-3, European Collection of Cell Cultures (ECACC) 90112714) was purchased from the ECACC (Salisbury, UK), and human negroid cervix epitheloid carcinoma (HeLa) cells were purchased from ATCC (Middlesex, UK). Cell culture reagents, namely Roswell Park Memorial Institute (RPMI) 1640 medium, Dulbecco's modified Eagle's medium (DMEM), trypsin/EDTA, and antibiotic–antimicotic solutions were obtained from Sigma (St Louis, MO, USA) while fetal bovine serum (FBS) was obtained in Biochrom (Berlin, Germany). In sample processing, TRI Reagent and the MEGAscript® Kit were acquired in Ambion (Carlsbad, CA, USA), and all chemicals used in the cell lysis buffer were from Sigma (St Louis, MO, USA). GreenSafe Direct Load stain for nucleic acid gel electrophoresis was purchased in NZYTech (Lisbon, Portugal). Arginine Sepharose 4B gel was obtained in Amersham Biosciences (Uppsala, Sweden). The compounds used in chromatographic experiments were sodium chloride purchased from Panreac (Barcelona, Spain) and Tris base from Fluka (Sigma). All solutions were freshly prepared using 0.01 % diethylpyrocarbonate (DEPC)-treated water from Fluka (Sigma), and the elution buffers were filtered through a 0.20-μm pore size membrane (Schleicher Schuell, Dassel, Germany) and degassed ultrasonically. The DNA molecular weight marker, HyperLadder I, was obtained in Bioline (London, UK). All experiments were performed with RNase-free disposables.

Cell culture

PC-3 or HeLa cells were initially cultured in 25 cm^2 flasks in RPMI 1640 or DMEM media, respectively at 37 °C in a humidified incubator in 95 % air/5 % CO_2. The culture media were supplemented with 10 % (*v*/*v*) FBS and 1 % (*v*/*v*)

antibiotic–antimicotic solution. The PC-3 or HeLa cells were routinely passaged at 90–95 % confluence to 75 cm^2 flasks.

Total RNA samples, cell lysates, and chemically synthesized RNA

To obtain total RNA samples and cell lysates, PC-3 or HeLa cells were collected after washing and detaching with 0.25 % trypsin/EDTA. The trypsin solution was neutralized with RPMI 1640 or DMEM supplemented culture medium and the cells were spun down in a centrifuge (Sigma 3K18C, Bioblock scientific) at 1,500 rpm for 5 min at room temperature, and resuspended in culture medium. The cells spin and resuspension was repeated to ensure complete removal of trypsin. The recovered cells were resuspended in 1 mL of PBS and counted using a Neubauer chamber. Subsequently, approximately 10×10^6cells were spun down and pellets were lysed according to the sample required. Thus, total RNA samples from PC-3 or HeLa cells were extracted and purified by TRI Reagent according to manufacturer's instructions while cell lysates were obtained as we have previously described [24]. Briefly, the pellets were resuspended in 2 mL denaturing cell lysis solution (4 M guanidinium thiocyanate; 25 mM sodium citrate, pH 4.0; 0.5 % (m/v) N-laurosylsarcosine; and 0.1 M β-mercaptoethanol) and after 5 min incubation at room temperature, 2 mL of ice-cold isopropanol were added. The lysate solution was homogenized by inversion and incubated on ice for 5 min. The precipitated molecules were recovered by centrifugation at 16,000×g for 20 min at 4 °C. The pellet was washed with 1 mL of 75 % ethanol in DEPC-treated water and incubated at room temperature for 10 min, followed by a 5-min centrifugation at 16,000×g (4 °C). The air-dried pellet was dissolved in 200 μL of 0.01 % DEPC-treated water and incubated 5 min at room temperature to ensure complete solubilization. Crude samples of chemically synthesized RNA were obtained by in vitro transcription using pTRI-Xef control template provided by the kit and following the manufacturer's instructions, but no further purification of RNA transcript was performed.

Total RNA samples from PC-3 cells were used as reference for the analytical method validation.

Agarose gel electrophoresis

All RNA-based samples were qualitatively assessed for purity and integrity by horizontal electrophoresis using 15-cm 1.0 % agarose gels (Hoefer, San Francisco, CA, USA), which were stained with GreenSafe Direct Load (0.25 μg/mL). Gel electrophoresis was carried out at 110 V with TAE buffer in DEPC-treated water (40 mM Tris base, 20 mM acetic acid, and 1 mM EDTA; pH 8.0). The nucleic acids were visualized under UV light in an UVIdoc system (Uvitec, Cambridge, UK). Next, the samples were applied in analytical arginine affinity chromatography to either perform method validation or test its applicability.

Chromatographic conditions

An ÄKTA Avant system controlled by UNICORN 6.1 software, both from GE Healthcare (Uppsala, Sweden), was used in the analytical chromatographic study. The separation of analytes was performed at room temperature on a Tricorn High-Performance Column (5/20) packed with 2 mL of the commercial Arginine–sepharose 4B gel, characterized by the manufacturer as a cross-linked 4 % beaded agarose matrix with a 12-atom spacer and an extent of labeling between 14 and 20 μmoL/mL. The system was initially prepared with ice-cold solutions of 10 mM Tris–HCl buffer, pH 8.0 in pump A and 1 M NaCl in 10 mM Tris–HCl buffer, pH 8.0 in pump B. The column was equilibrated with a mobile phase consisting of 21 % of buffer B at a flow rate of 0.5 mL/min. Conductivity was rigorously controlled at 19 mS/cm. RNA standards or other RNA-based samples were injected using a 50-μL loop using the same conditions (flow rate, buffer percentage, and conductivity). After a 6-min run, the ionic strength of the mobile phase was increased to 100 % buffer B and the flow rate decreased to 0.3 mL/min. These conditions were kept for 15 min. The chromatographic runs were monitored by an UV detector at 260 nm. To identify the eluting species, peak fractions were pooled according to the chromatograms obtained and comprising the full amplitude of the peak. After each run, the column was washed with DEPC-treated water and re-equilibrated with equilibration buffer for 10 min.

Results and Discussion

Recent findings concerning RNA isolation demonstrated the applicability of arginine affinity chromatography in eukaryotic RNA purification, as confirmed by gene expression analysis [24]. Therefore, in the present work the analytical method was developed based on the binding and elution conditions previously described for preparative arginine-based chromatography. Briefly, a NaCl concentration of approximately 210 mM can be used to retain RNA molecules and eliminate impurities and by increasing the ionic strength to 1 M NaCl, bound RNA is recovered [24]. Selectivity, linearity, accuracy, lower quantification and detection limits, reproducibility, and repeatability were evaluated as these are the essential characteristics of a bioanalytical method to ensure the reliability of an analytical result, according to Food and Drug Administration [29] and European Medicines Agency (EMA) [30] guidelines.

Selectivity

Selectivity is an extremely relevant parameter as it is required for the implementation of a new analytical method able to recognize and quantify RNA molecules. Actually, specific RNA recognition was one of the main capacities found in arginine-based affinity chromatography in our previous work [24]. Overall, it was suggested that electrostatic interactions were playing an important role on RNA retention due to the predominant negative charge of the phosphate groups and the positive charge of arginine amino acid, at the pH in study. Additionally, the nucleotide-based contacts with arginine side chain were also involved and modulate the interaction and specificity found in arginine–agarose chromatography. Figure 1 illustrates the suggested interactions.

Therefore, the ability of the proposed bioanalytical method to measure and differentiate RNA in the presence of impurities was confirmed by analyzing the elution profile and retention times of PC-3 cell lysate samples containing gDNA, total RNA and proteins, against PC-3 total RNA samples used as reference. Nonetheless, purity and integrity of the samples were qualitatively evaluated by electrophoresis before the chromatographic runs. Figure 2 (lane R) shows the electrophoretic profile of a total RNA sample, including 28S rRNA, 18S rRNA, mRNA, and small RNAs, as 5S RNA and regulatory non-coding RNAs, such as siRNA and miRNA; in lane L, it is presented the cell lysate profile which besides RNA species also shows a high molecular weight band corresponding to gDNA.

Figure 3a, b shows the chromatograms obtained after injection of each sample onto the analytical arginine column. The sample containing total RNA (Fig. 3a) displays a very small peak in the first gradient step and a distinguished second peak at 1 M NaCl in Tris buffer step, with a retention time of 15.46 min. The RNA peak is often broad and tailing because of the different RNA species eluting at this experimental condition. For RNA standards, this last peak was the only expected, however, chromatographic analysis showed the presence of another peak, which was found to be related to a persistent gDNA contamination that was not detected under quality control electrophoresis analysis (Fig. 2). In fact, the manufacturer's instructions of TRI Reagent forewarn for a residual presence of DNA and recommends treating RNA samples with DNase for subsequent applications. In this study, the analytical arginine affinity chromatography was able to detect gDNA contamination and remove it completely from total RNA samples. An evidence of this ability is shown in Fig. 3b which refers to the injection of a complex lysate

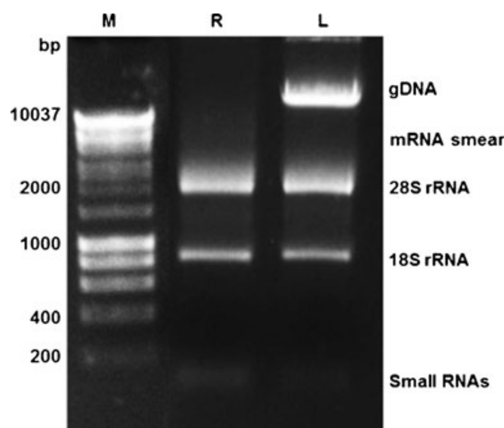

Fig. 2 Agarose gel electrophoresis of reference RNA (*lane R*) and cell lysates (*lane L*) from PC-3 cells. *Lane M*, DNA molecular weight marker

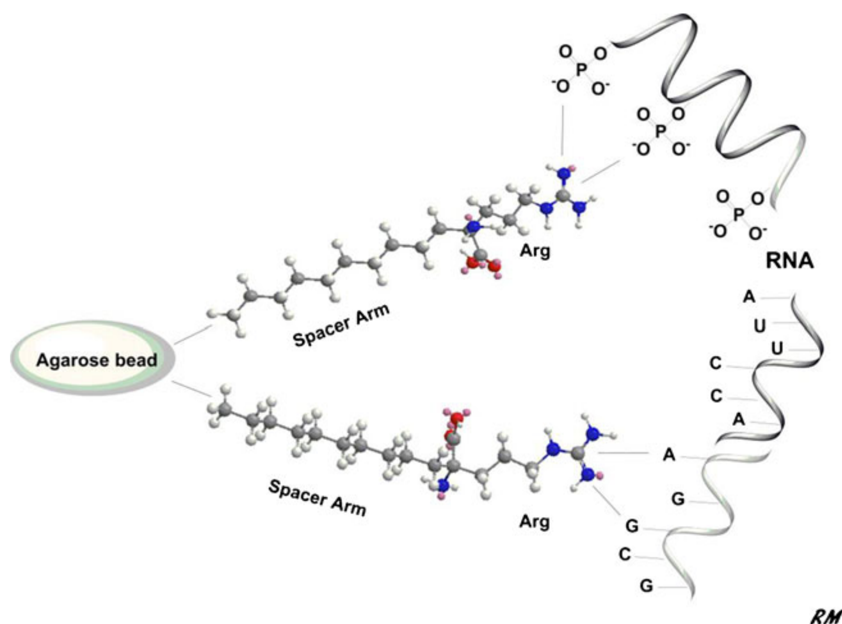

Fig. 1 Proposed interactions between arginine–agarose matrix and RNA molecules. Given the multiplicity of arginine side-chain interactions and depending upon RNA folding state, arginine will preferably bind to phosphate groups of RNA backbone or RNA bases

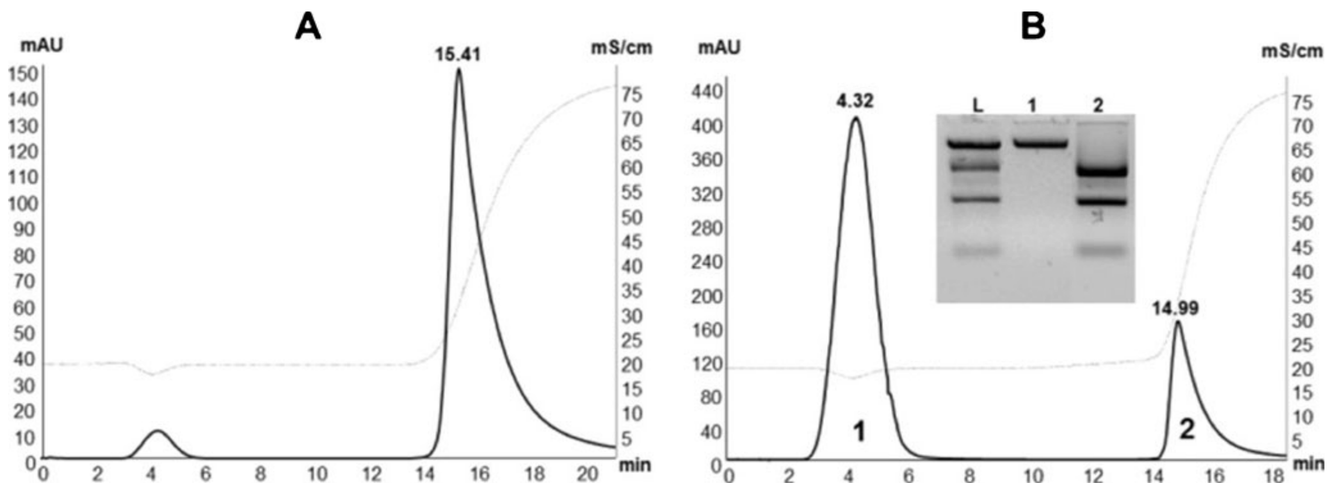

Fig. 3 Chromatographic analysis of RNA using the analytical method based on arginine affinity chromatography. (**A**) Chromatographic profile of standard RNA. (**B**) Chromatographic profile of cell lysates and electrophoresis image of nucleic acids content in each peak. Fractions corresponding to peaks (*1*) and (*2*) are shown in *lanes 1* and *2*, respectively. *Lane L*, PC-3 cell lysate injected onto the analytical arginine column

sample. Thus, a first peak was obtained in the first gradient step at 4.32 min., and a second peak was achieved after increasing the NaCl concentration to 1 M, with a retention time of 14.99 min., being this last peak coincident with the RNA peak obtained in chromatographic analysis of total RNA (Fig. 3a). To ensure nucleic acids content of each peak, fractions from the full amplitude of the peaks were collected and treated according to Martins and co-workers [24]. The electrophoretic analysis presented in Fig. 3b shows the presence of gDNA in the first peak (lane 1; peak 1) while RNA molecules were isolated in the second peak (lane 2; peak 2). Thus, these experiments confirmed the selectivity of the analytical method to effectively distinguish RNA molecules between other biomolecules present in cell lysates. Moreover, despite the fact of gDNA and RNA have similar physicochemical properties; a good resolution between the two peaks was achieved, which supports the specificity of arginine-based chromatography.

Discrimination between RNA and DNA is a great advantage for this method because DNA is one of the major interfering products in RNA-based analysis [31] and many purification methods do not fully remove it from RNA preparations. In addition, common quantification methods, as spectrophotometric analysis or intercalating fluorescence dye-based systems, do not have the sensibility to distinguish them as they absorb at 260 nm, thereby, false quantification readings, and irreproducible data may often occur [32].

Furthermore, the high selectivity of arginine matrix for RNA molecules in complex mixtures can be view as an advantage in clinical investigations, such as the detection and quantification of potential biomarkers like microRNAs (miRNAs). There is presently a lack of consensus regarding optimal methodologies for miRNA quantification [33]. Today, miRNAs quantification are preferably performed by RT-qPCR because of its substantial sensitivity and specificity [34]; however, this technique follows several initial steps of sample collection and preparation, becoming very time consuming and expensive. The proposed method presents the advantage of detecting and quantifying RNA molecules in samples that just require the disruption of cells without the need for specific extraction procedures. This will simplify the process and improve economics.

Linearity, accuracy, and detection and quantification limits

The linearity of an analytical procedure is defined by EMA "as its ability (within a given range) to obtain test results which are directly proportional to the concentration (amount) of analyte in the sample" [30]. Thus, different concentrations of RNA in a range of 20–200 ng/μL (20, 30, 50, 70, 100, 120, 150, and 200 ng/μL) were tested by the analytical method to verify a linear relationship. RNA standards were prepared in DEPC-treated water by dilution from a stock solution of total RNA isolated by the commercial TRI reagent. The initial concentration of the stock solution was assessed through spectrophotometric measurement using a NanoPhotometer (IMPLEN, Munich, Germany). Within the range studied, the method was found linear considering either height or area of the RNA peak. Statistical analysis demonstrated a good correlation coefficient of 0.9991. Figure 4 presents the regression data obtained from the correlation between RNA peak area and concentration. In addition, an example of the chromatographic profiles of three RNA standards is shown in Fig. 5, demonstrating the consistency of RNA elution at 15.45 min. Nonetheless, the peaks of the different nucleic acids show low correlation between experiments due to the heterogeneous content of the initial cell lysates, which presents unrelated gDNA concentrations.

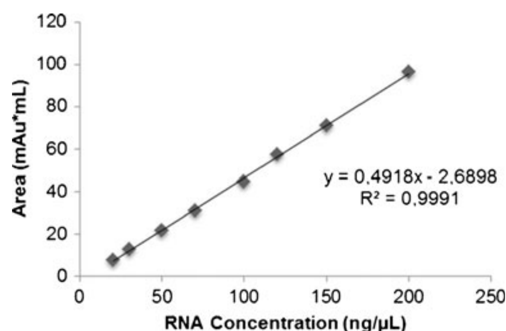

Fig. 4 Linear regression data obtained from the plot of peak areas as function of RNA concentration. R^2 coefficient of determination of the linear regression

Table 1 shows the results of accuracy assessed by back-calculated concentration for all standards and for an intermediate RNA concentration of 80 ng/μL. Thus, the closeness agreement between the values of the nominal concentrations and RNA concentrations calculated using the regression equation was evaluated by determining the coefficient of variation (CV) and the mean relative error. Both parameters were within the range of ±15 % of the nominal concentration for all RNA concentrations studied, including the lowest concentration standard of 20 ng/μL which shows a relative error of 5.24 % (Table 1), which is far below the value of 20 % accepted by the guidelines [30]. In fact, the lowest RNA standard was established as the limit of quantification, that was defined as the minimal RNA amount for which the arginine-based chromatography method has a suitable level of precision (a CV lower than ±20 %) and accuracy (within ±20 % of the nominal concentration). Additionally, the 20 ng/μL was also considered the limit of detection taking into account the baseline noise. This parameter was determined by comparing chromatographic signals of blank samples containing 1 M NaCl in Tris buffer pH 8 and the lowest RNA concentration. The assays showed a signal-to-noise ratio of 2:1 which is generally acceptable for estimating the detection limit [30]. Although common methods for RNA quantification show lower limits of quantification than the proposed method, their range of quantification is also low, normally between 3 μg/μL and 2 ng/μL for nano-spectrophometers or 500–0.05 ng/μL, for bioanalyzers [35]. This can limit the quantification of RNA produced in a large scale, as in the case of RNA vaccines [36]. Therefore, the proposed method presents an advantage over other quantification methods since it offers the possibility to increase the upper limit of the quantification range validated here. Based on our previous work, arginine affinity chromatography allows the quantification of at least 35 μg of RNA [24] nevertheless the determination of dynamic binding capacity of this matrix is required to establish the highest quantification limit.

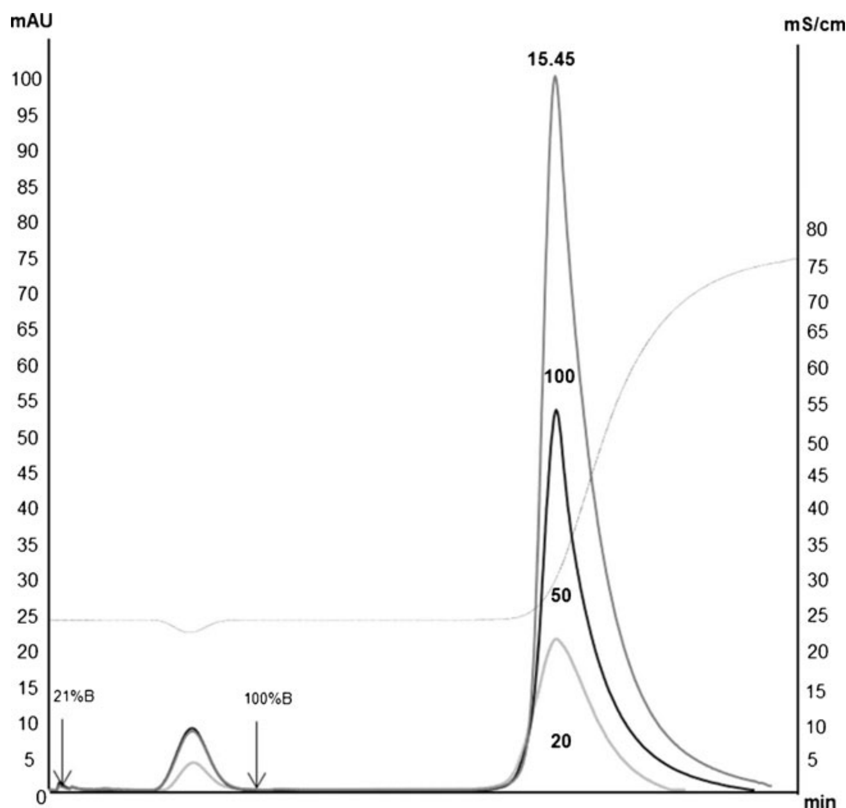

Fig. 5 Representative chromatograms of three RNA standards with concentrations of 20, 50, and 100 ng/μL used in the validation of the analytical method

Table 1 Evaluation of accuracy of the proposed analytical method ($n=3$)

[Nominal] (ng/μL)	Mean (RNA)±SD (ng/μL)	CV (%)	Mean relative error (%)
20	20.63±0.39	5.24	0.03
30	32.98±0.75	5.55	0.10
50	50.02±0.71	3.22	0.00
70	68.28±0.28	0.92	−0.02
100	97.89±0.47	1.03	−0.02
120	121.96±0.29	0.51	0.02
150	148.58±0.08	0.11	−0.10
200	204.81±0.35	0.36	0.02
80	81.48±0.57	1.52	0.02

SD standard deviation, *CV* coefficient of variation

Precision

The investigation of precision reflects the closeness agreement of individual measurements of RNA when the analytical procedure is applied repeatedly to multiple sampling of homogeneous RNA reference material, and it can be studied at two levels: repeatability and reproducibility, also termed as intra- and inter-day precision, respectively. The values for intra-day precision were obtained from consecutive injections of standard RNA of a minimum of three concentration levels (low, medium, and high) over a minimal of nine determinations. The same samples were also analyzed over a period of three consecutive days to determine the inter-day precision. The series of measurements were analyzed by determining statistical parameters related to precision which are presented in Tables 2 and 3. The calculated mean, standard deviation, CV, and mean relative error for three RNA concentrations of 30, 80, and 150 ng/μL concerning repeatability (Table 2) and reproducibility (Table 3) showed an adequate precision for this new method. The CV values were lower than 2 or 3 % in intra- and inter-day precision, respectively. Moreover, the deviation from the nominal concentration (mean relative error) was between −1.17 and 3.59 % for intra-day experiments and inter-day values did not exceed 5.33 %, which also confirms the accuracy found for the proposed analytical method.

Building on the successful results present so far, the method is concluded to be simple, selective, linear, precise, and accurate in a concentration range adequate to measure RNA content and to assess its purity. Moreover, several flaws in current RNA quantification techniques may be overcome using the described analytical arginine affinity method.

Method application

The applicability and robustness of the analytical arginine-affinity chromatography method for RNA quantification was further explored by using RNA samples from different sources and preparation methods. The intention of these assays was to evaluate the broad use of the method regardless the origin of the RNA or its purpose. Therefore, total RNA obtained from HeLa cells and crude samples of RNA chemically synthesized by in vitro transcription were tested according to the procedure previously described for RNA-based samples from PC-3 cells. Figure 6 shows the chromatographic analysis of total RNA samples from HeLa cells with an electrophoretic image of the injected sample (lane H). The displayed peaks show similar retention times when compared with the chromatographic profiles of total RNA from PC-3 cells (Fig. 3). In case of peak height, RNA from HeLa cells revealed lower peaks but broader than PC-3 samples. The differences in chromatographic signals were found to be due to the predominance of different species of RNA in the samples. Accordingly, electrophoretic analysis of RNA molecules isolated from each cell line also allowed the virtual perception of lower molecular

Table 2 Determination of intra-day variability of the analytical method for RNA assessment ($n=3$)

Intra-day precision			
[Nominal] (ng/μL)	Mean (RNA)±SD (ng/μL)	CV (%)	Mean relative error (%)
30	31.10±0.30	0.98	3.59
80	81.59±1.18	1.45	1.99
150	148.24±0.86	0.58	−1.17

SD standard deviation, *CV* coefficient of variation

Table 3 Determination of inter-day variability of the analytical method for RNA assessment ($n=3$)

Inter-day precision			
[Nominal] (ng/μL)	Mean (RNA)±SD (ng/μL)	CV (%)	Mean relative error (%)
30	31.60±0.94	2.97	5.33
80	81.73±0.39	0.48	2.16
150	148.00±0.41	0.27	−1.33

SD standard deviation, *CV* coefficient of variation

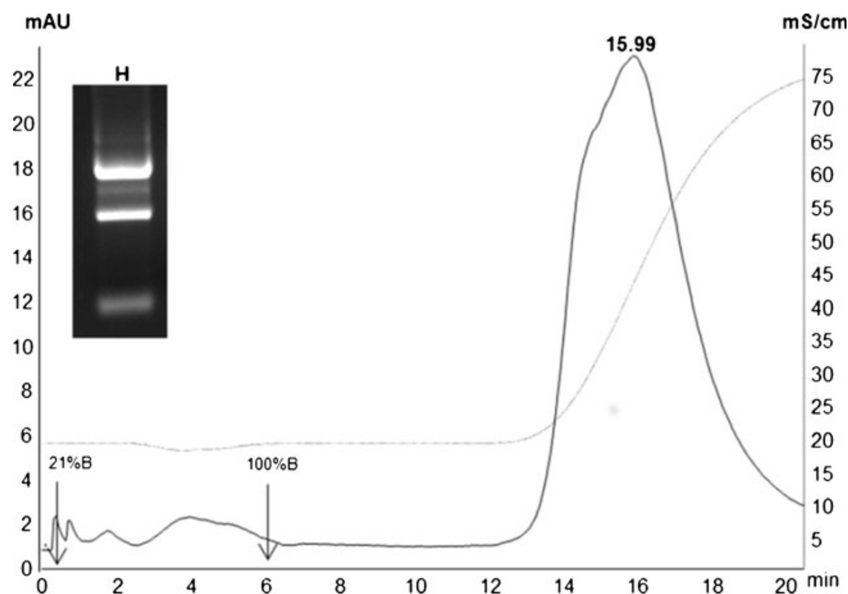

Fig. 6 Chromatographic analysis of RNA isolated from HeLa cells and electrophoretic profile of the RNA sample (*lane H*) injected onto the analytical arginine column

weight RNA species in HeLa cells (Fig. 6) than in PC-3 cells (Fig. 2). The isolation of dissimilar RNA species from each cell line could be expected because of their unparalleled life cycles, which are differently favoring the production of one type of RNA over another to accomplish physiological cell function [37]. Nevertheless, peak areas of the injected concentrations of RNA from HeLa cells are consistent with RNA standards from PC-3 cells, giving concordant results between the injected concentration and the concentration achieved by extrapolation of the peak area on the regression equation.

With regard to the applicability of the proposed analytical method to monitor and quantify chemically synthesized RNA from a crude sample obtained by in vitro transcription, arginine column also demonstrated an appropriate performance for this purpose. In these assays, the applicability of the method was attested either by the evaluation of the chromatographic behavior or by the electrophoretic analysis of the eluting peaks. Figure 7 shows the chromatographic profile of a crude sample of chemically synthesized RNA together with the electrophoretic image of the injected sample and the resulting peaks. Again, the second peak corresponds to RNA elution, confirmed by the similar retention times and electrophoretic analysis (peak 2; lane 2). However, the first peak presented a retention time of 4.94 min., which is not coincident with the retention time of gDNA from PC-3 cell lysates (4.32 min). Based on the electrophoretic result, it was hypothesized that the peak did not refer to the elution of a nucleic acid. Thus, as no previous purification was performed in the transcription-based RNA samples, the existence of a first peak can be due to impurities, such as enzymes, nucleotides, salts,

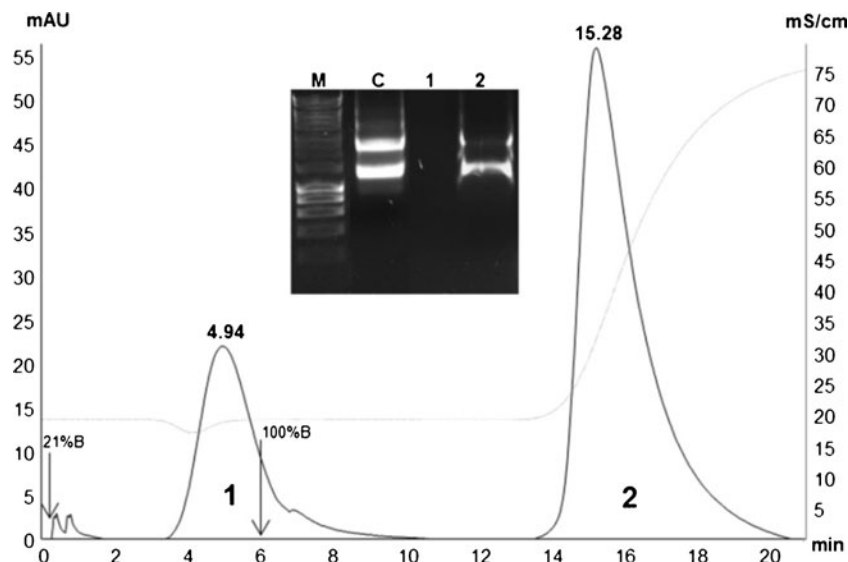

Fig. 7 Chromatographic profile of crude sample of chemically synthesized RNA and electrophoretic analysis of the injected sample and resulting peaks. Fractions corresponding to peaks (*1*) and (*2*) are shown in *lanes 1* and *2*, respectively. *Lane C*, crude sample of chemically synthesized RNA by in vitro transcription injected onto the analytical arginine column

buffer or other components that derive from the in vitro transcription process and may absorb at 260 nm. Therefore, the accurate quantification of chemically synthesized RNA can also be achieved by arginine affinity chromatography method, nonetheless further evaluation of the presence of impurities in RNA peaks should be assessed by complementary techniques.

Conclusions

The requirement of new analytical techniques for rigorous analysis and quantification of RNA samples meeting regulatory demands is becoming evident as basic research improves the understanding of biological roles of RNA and pharmaceutical industry starts to focus on it as a biotherapeutic tool.

In this work, an analytical technique based on arginine-affinity chromatography for quantification and quality verification of total RNA was developed and validated. The method was found linear in the range of 20–200 ng/μL, precise and accurate, according to international guidelines. The specificity was determined with cell lysates, showing that gDNA is well separated from RNA molecules, which confirms the great selectivity of arginine column and highlights the great advantage of this method over other quantification techniques. Moreover, the current methods used for RNA detection and monitorization were discussed and compared with the proposed analytical technique, underlining the several advantages of the new methodology.

Furthermore, the proposed method was successfully employed in the quantification and monitorization of different RNA types, namely RNA isolated from HeLa cells and chemically synthesized by in vitro transcription, which demonstrates its potential applicability in basic research and RNA-based therapies.

Acknowledgments This work was supported by FCT, the Portuguese Foundation for Science and Technology (PTDC/EBB-BIO/114320/2009 and PEst-C/SAU/UI0709/2011 COMPETE). Rita Martins also acknowledges a fellowship (SFRH/BD/ 64100/2009) from FCT.

Conflict of interest The authors declare no conflict of interest.

References

1. Sharp PA (2009) The centrality of RNA. Cell 136(4):577–580
2. Kreiter S, Diken M, Selmi A, Tureci O, Sahin U (2011) Tumor vaccination using messenger RNA: prospects of a future therapy. Curr Opin Immunol 23(3):399–406
3. Vermeulen J, De Preter K, Lefever S, Nuytens J, De Vloed F, Derveaux S, Hellemans J, Speleman F, Vandesompele J (2011) Measurable impact of RNA quality on gene expression results from quantitative PCR. Nucleic Acids Res 39(9):e63
4. Gjerde DT, Hoang L, Hornby D (2009) RNA purification and analysis: sample preparation, extraction, chromatography. Wiley, Weinheim
5. Farrell RE (2012) RNA methodologies. In: Meyers RA (ed) Epigenetic regulation and epigenomics, vol 2, 1st. Wiley, Weinheim, pp 1–37
6. Wieczorek D, Delauriere L, T. S (October 2012) Methods of RNA quality assessment. Promega Corporation Website. http://worldwide.promega.com/resources/articles/pubhub/methods-of-rna-quality-assessment. Accessed 20 Feb 2013
7. Heptinstall J, Rapley R (2000) Spectrophotometric analysis of nucleic acids. In: Rapley R (ed) The nucleic acid protocols handbook. Humana Press, Totowa, pp 57–60. doi:10.1385/1-59259-038-1:57
8. Barril P, Nates S (2012) Introduction to agarose and polyacrylamide gel electrophoresis matrices with respect to their detection sensitivities. In: Magdeldin DS (ed) Gel electrophoresis—principles and basics. InTech, Rijeka. doi:10.5772/38573
9. Ohta T, Tokishita S, Yamagata H (2001) Ethidium bromide and SYBR Green I enhance the genotoxicity of UV-irradiation and chemical mutagens in *E. coli*. Mutat Res 492(1–2):91–97
10. Jones LJ, Yue ST, Cheung C-Y, Singer VL (1998) RNA quantitation by fluorescence-based solution assay: RiboGreen reagent characterization. Anal Biochem 265(2):368–374
11. Schmittgen TD, Lee EJ, Jiang J, Sarkar A, Yang L, Elton TS, Chen C (2008) Real-time PCR quantification of precursor and mature microRNA. Methods 44(1):31–38
12. Masotti A, Preckel T (2006) Analysis of small RNAs with the Agilent 2100 Bioanalyzer. Nat Methods. doi:10.1038/NMETH908
13. Dickman MJ (2011) Ion pair reverse-phase chromatography: a versatile platform for the analysis of RNA. Chromatography Today 4:22
14. McCarthy SM, Gilar M, Gebler J (2009) Reversed-phase ion-pair liquid chromatography analysis and purification of small interfering RNA. Anal Biochem 390(2):181–188
15. Azarani A, Hecker KH (2001) RNA analysis by ion-pair reversed-phase high performance liquid chromatography. Nucleic Acids Res 29(2):E7
16. Dickman MJ, Hornby DP (2006) Enrichment and analysis of RNA centered on ion pair reverse phase methodology. RNA 12(4):691–696
17. Waghmare SP, Pousinis P, Hornby DP, Dickman MJ (2009) Studying the mechanism of RNA separations using RNA chromatography and its application in the analysis of ribosomal RNA and RNA:RNA interactions. J Chromatogr A 1216(9):1377–1382
18. Clonis YD (2006) Affinity chromatography matures as bioinformatic and combinatorial tools develop. J Chromatogr A 1101(1–2):1–24
19. Magdeldin S, Moser A (2012) Affinity chromatography. In: Magdeldin DS (ed) Affinity chromatography: principles and applications. InTech, Rijeka. doi:10.5772/39087
20. Srisawat C, Goldstein IJ, Engelke DR (2001) Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures. Nucleic Acids Res 29(2):E4
21. Batey RT, Kieft JS (2007) Improved native affinity purification of RNA. RNA 13(8):1384–1389
22. Martins R, Queiroz JA, Sousa F (2010) A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography. J Mol Recognit 23(6):519–524
23. Martins R, Queiroz JA, Sousa F (2012) Histidine affinity chromatography-based methodology for the simultaneous isolation of *Escherichia coli* small and ribosomal RNA. Biomed Chromatogr 26(7):781–788
24. Martins R, Maia CJ, Queiroz JA, Sousa F (2012) A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography. J Sep Sci 35(22):3217–3226
25. Sousa F, Cruz C, Queiroz JA (2010) Amino acids–nucleotides biomolecular recognition: from biological occurrence to affinity chromatography. J Mol Recognit 23(6):505–518

26. Treger M, Westhof E (2001) Statistical analysis of atomic contacts at RNA–protein interfaces. J Mol Recognit 14(4):199–214
27. Jeong E, Kim H, Lee SW, Han K (2003) Discovering the interaction propensities of amino acids and nucleotides from protein–RNA complexes. Mol Cells 16:161–167
28. Yarus M, Widmann JJ, Knight R (2009) RNA–amino acid binding: a stereochemical era for the genetic code. J Mol Evol 69(5):406–429
29. U.S. Department of Health and Human Services (2001) Guidance for industry, bioanalytical method validation. UD Food and Drug Administration, Rockville
30. Kaighn ME, Narayan KS, Ohnuki Y, Lechner JF, Jones LW (1979) Establishment and characterization of a human prostatic carcinoma cell line (PC-3). Investig Urol 17(1):16–23
31. Laurell H, Iacovoni JS, Abot A, Svec D, Maoret JJ, Arnal JF, Kubista M (2012) Correction of RT-qPCR data for genomic DNA-derived signals with ValidPrime. Nucleic Acids Res 40(7):e51
32. Fleige S, Pfaffl MW (2006) RNA integrity and the effect on the real-time qRT-PCR performance. Mol Aspects Med 27(2–3):126–139
33. Mo MH, Chen L, Yebo F, Wang W, Fu SW (2012) Cell-free circulating miRNA biomarkers in cancer. J Cancer 3:432–448
34. Kang K, Peng X, Luo J, Gou D (2012) Identification of circulating miRNA biomarkers based on global quantitative real-time PCR profiling. J Anim Sci Biotechnol 3(1):4
35. Aranda R, Dineen SM, Craig RL, Guerrieri RA, Robertson JM (2009) Comparison and evaluation of RNA quantification methods using viral, prokaryotic, and eukaryotic RNA over a 10(4) concentration range. Anal Biochem 387(1):122–127
36. Pascolo S (2006) Vaccination with messenger RNA. Methods Mol Med 127:23–40
37. Salem C, El-Alfy M, Leblond CP (1998) Changes in the rate of RNA synthesis during the cell cycle. Anat Rec 250(1):6–12

Paper VI

Arginine-affinity chromatography for mRNA vaccines purification

R. Martins, C. J. Maia, J. A. Queiroz, F. Sousa

(to submit)

Arginine-affinity chromatography for mRNA vaccines purification

R. Martins, C. J. Maia, J. A. Queiroz, F. Sousa*

CICS-UBI – Health Sciences Research Centre, University of Beira Interior, Av. Infante D. Henrique, 6200-506 Covilhã, Portugal.

*corresponding author:
E-mail: fani.sousa@fcsaude.ubi.pt
Telephone: +351 275 329 002
Fax: +351 275 329 099

Abstract

Coding messenger RNA (mRNA) is emerging as a particularly attractive option in the development of new approaches for the treatment of cancer or infection diseases focusing on immunotherapies. Cervical cancer is the second most common cancer in women worldwide. Human papillomavirus (HPV) is the primary etiologic agent of cervical cancer and the two HPV16 proteins, E6 and E7, are prospective therapeutic vaccines as they are consistently expressed in tumour cells.

Because of the increasing success of mRNA-based immunotherapies, the development of new tools to improve RNA purification is of great significance since rigorous quality criteria recommended by international regulatory authorities should be fulfilled. Therefore, a recently described chromatographic strategy for RNA purification was considered, exploiting affinity interactions between immobilized amino acids and nucleic acids.

In the present study, a single arginine affinity chromatography step was employed for the purification of mRNA, produced by enzymatic synthesis, encoding HPV16 E6 and E7 proteins. The chromatographic experiments were developed by optimizing a sodium chloride elution gradient to purify the mRNA transcripts from impurities resultant from the production process. The evaluation of the purification process revealed that both E6 and E7 mRNAs are recovered with high yield, 97.5% and 98%, respectively. Furthermore, the quality control showed that mRNA preparations presented high integrity as well as an improved purity. The results obtained so far indicated that the chromatographic separation is a promising strategy for mRNA purification pursuing therapeutic applications.

Keywords

Affinity chromatography, arginine, messenger RNA, *in vitro* transcription.

1. Introduction

Recently, messenger RNA (mRNA) is being an important target of several research studies, due to its potential applicability in the treatment of cancer or infection diseases focusing on immunotherapies (Kreiter *et al.*, 2011). The physiological role of mRNA is to transfer genetic information from the nucleus to the cytoplasm where this information is translated into the corresponding protein. The concept of mRNA vaccination is to carry the information of an antigenic protein to be translated in the cell cytoplasm and therefore generate an immunological response (Bringmann *et al.*, 2010). Recent studies describe that mRNAs encoding antigens are able to stimulate all effectors of the adaptive immune response: B lymphocytes, cytotoxic T cells, and T helper cells (Ponsaerts *et al.*, 2003). Besides, it has demonstrated several advantages to be applied as a vaccine including feasibility, applicability, safeness, and effectiveness when it comes to the generation of immune responses (Pascolo, 2008, Van Lint *et al.*, 2013). In particular, the safety of mRNA-based treatments supports the use of mRNA-vaccination for therapeutic or prophylactic approaches as it addresses some shortcomings of recombinant virus or DNA-based vaccination therapies (Vergati *et al.*, 2010).

Cervical cancer is the second most common cancer in women worldwide. Human papillomavirus (HPV) is the primary etiologic agent of cervical cancer and HPV16 is by far the most common carcinogenic type (Huh, 2001). Two HPV16 proteins, E6 and E7, are consistently expressed in tumour cells and are critical to the induction and maintenance of cellular transformation. Thus, E6 and E7 proteins represent good targets for developing therapeutic vaccines for cervical cancer (Bharti *et al.*, 2009). Currently, a prophylactic vaccine of HPV16 and HPV18 has been registered with successful results in cervical cancer prevention, but it has no benefit in women who were already infected with those HPV types (Franco *et al.*, 2009). Thus, the necessity to develop effective therapeutic HPV vaccines continues and the use of mRNA-based vaccination has been described to be a promising approach (Thornburg *et al.*, 2000).

Due to the increasing success of clinical application of mRNA-based immunotherapies, the development of new tools to improve RNA manufacture process is of great significance. Of particular importance is the purification process, since the RNA bioproduct should fulfill rigorous quality criteria recommended by international Regulatory Authorities, such as the World Health Organization (WHO), Food and Drug Administration (FDA) and European Medicines Agency (EMA) (Pascolo, 2008).

Therapeutic mRNA is currently produced by cloning the gene encoding a desired protein in a plasmid containing regulatory sequences (poly(A) tail, untranslated regions (UTRs)), followed by amplification in *Escherichia coli* (*E. coli*). Next, the plasmid is recovered, linearized,

purified and *in vitro* transcribed using a RNA polymerase that recognizes the bacteriophage promoter (Weide *et al.*, 2008). However, final RNA products need to be isolated from impurities derived from the synthesis process. In general, these impurities are the plasmid DNA (pDNA) template, proteins, nucleotides and aberrant oligonucleotides, salts and buffers. The current purification strategies for synthetic mRNA include the precipitation with high concentration of lithium chloride (LiCl) (Pascolo, 2006). However, most of the time this precipitation does not completely remove traces of contaminating pDNA or proteins and worse it introduces lithium metal into preparations. Thus, to improve mRNA purity, phenol/chloroform extraction often accompanies LiCl precipitation (Pascolo, 2006). Although, this methodology achieves highly pure mRNA preparations, it employs extremely hazardous compounds, which are not recommended to be used in clinical applications. Otherwise, mRNA can be purified using affinity chromatography and, in fact, this is strongly recommended for therapeutic applications of mRNA, as in vaccination approaches (Pascolo, 2008). However, the description of these affinity strategies is not totally available in the literature.

Therefore, new chromatographic strategies for mRNA purification were considered, exploiting affinity interactions between amino acids and nucleic acids. The potential of using affinity chromatography with histidine or arginine as ligands has been recently demonstrated by our research group to selectively isolate and monitor different RNA species of various cell sources (Martins *et al.*, 2012a, Martins *et al.*, 2010, Martins *et al.*, 2012b, Martins *et al.*, 2013). Histidine matrix demonstrated to mediate a specific recognition for a regulatory non-coding RNA (ncRNA), 6S RNA, allowing its purification from a complex mixture of other *E.coli* ncRNAs (Martins *et al.*, 2010). Furthermore, the simultaneous isolation of ncRNA and ribosomal RNA from *E.coli* cell lysates, eliminating host DNA and proteins, was also attained by using an histidine chromatography-based method (Martins *et al.*, 2012b). Moreover, arginine-agarose support enabled the isolation of total RNA from impurities of an eukaryotic cell extract (Martins *et al.*, 2012a) and was also used in the development of an analytical methodology for quantification and monitorization of RNA molecules (Martins *et al.*, 2013). Hence, this work intends to explore the possibility of using arginine affinity chromatography to improve the purification of mRNA transcripts encoding the oncogenic proteins of HPV16 pursuing the development of a new and enhanced methodology for mRNA based-vaccination.

2. Experimental section

2.1. Materials

For plasmid design, 1321 HPV16 E6/E7 and T7TS plasmids were obtained in Addgene, a plasmid repository (Cambridge, USA). Synthetic oligonucleotides used as primers were purchased in StabVida (Lisboa, Portugal). Restriction enzymes BglII and EcoRV were acquired in Fermentas (Thermo fisher scientific Inc) and T4 DNA ligase was from Promega (Madison, USA). For production of mRNA molecules, MEGAscript® Kit was acquired in Ambion (Carlsbad, CA, USA) and Sma1 restriction enzyme was from Takara Bio Inc. (Shiga, Japan).

For the purification process, Arginine Sepharose 4B gel was obtained in GE Healthcare (Uppsala, Sweden). The compounds used in chromatographic experiments were sodium chloride (NaCl) purchased from Panreac (Barcelona, Spain) and Tris base from Fluka (Sigma, St Louis, MO, USA). All solutions were freshly prepared using 0.01% diethylpyrocarbonate (DEPC)-treated water from Fluka (Sigma) and buffers were filtered through a 0.20 µm pore size membrane (Schleicher Schuell, Dassel, Germany) and degassed ultrasonically. For the nucleic acid gel electrophoresis it was used GreenSafe Direct Load stain from NZYTech (Lisbon, Portugal) and DNA molecular weight marker, HyperLadder I, was obtained in Bioline (London, UK). All experiments were performed with RNase-free disposables.

2.2. Plasmid construct

HPV16 E6 and E7 genes were isolated by polymerase chain reaction (PCR) from the 1321 HPV-16 E6/E7 plasmid (Addgene plasmid 8641 (Munger *et al.*, 1989)), which contains the E6 and E7 genes. A set of primers corresponding to the 5' and 3' regions of E6 or E7 genes were used (Table 1). PCR was performed under standard conditions (95ºC for 30 sec, 60ºC for 30 sec and 72ºC for 15 sec, for 35 cycles) and the products were separated on a 1% agarose gel electrophoresis. The band corresponding to the expected PCR product size was cut out and the DNA was purified by NucleoSpin® Gel and PCR clean-up (Macherey-Nagel GmbH & Co, Düren, Germany). Purified DNA was cut with BglII and EcoRV and ligated using T4 DNA ligase into a similarly cut prokaryotic expression vector T7TS plasmid (Addgene 17091) (figure 1). The ligation mixture was used to transform DH5α *E. coli* competent cells by heat shock and selected for ampicillin resistance. Isolated colonies were collected to check for the presence of plasmid. The cloned HPV16 E6 or E7 genes in pT7TS were verified by PCR after performing the two restriction enzymes digestion (figure 2). E6 and E7 coding regions were verified by sequencing (StabVida, Lisboa, Portugal). Finally, the pT7TS E6 and pT7TS E7 (figure 1) were amplified by fermentation of *E. coli* DH5α in shake flasks at 37ºC and 250 rpm with 250 mL of Terrific Broth medium (12 g/L tryptone, 24 g/L yeast extract, 4 mL/L glycerol, 0.017 M KH_2PO_4, 0.072 M K_2HPO_4). Cell growth was suspended in the late log phase (OD600 ≈ 12). Cells

were recovered by centrifugation and the plasmids were extracted by NZYMaxiprep (NZYTech, Lisboa, Portugal) and stored at -20ºC or readily used in the production of mRNA molecules.

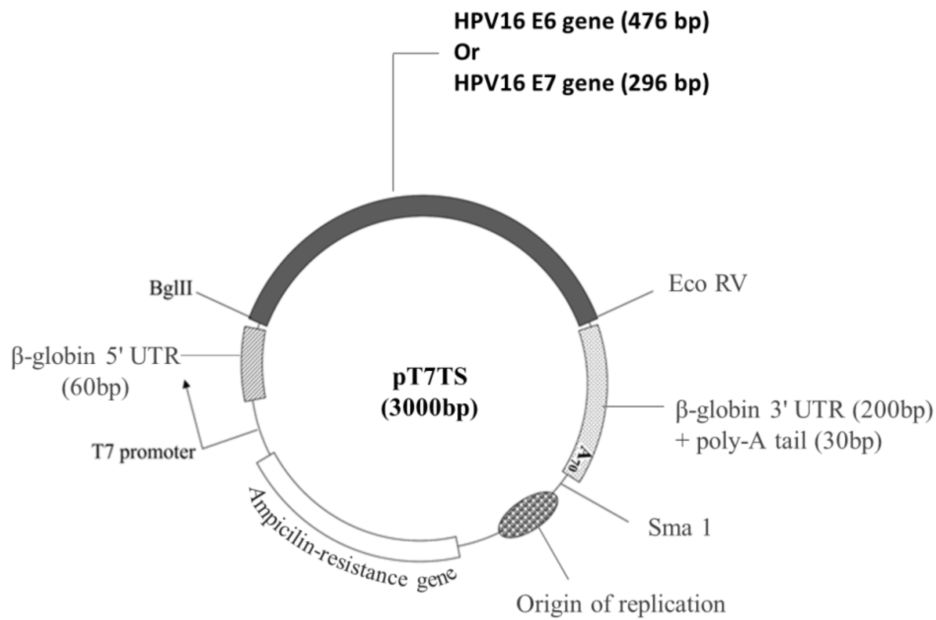

Figure 1 - Graphic map of the pT7TS plasmid backbone with the inserted fragments HPV16 E6 or HPV16 E7 genes. The restriction sites used for cloning E6 or E7 genes (BglII and EcoRV) and the restriction site used for linerarization (Sma1) are shown. The hatched box represent the β-globin 5' UTR (approx. 60 bp); the gray box represents the place where HPV16 E6 or HPV16 E7 coding genes (E6, approx 476bp and E7, approx 296bp) are inserted; the dotted box represents the β-globin 3' UTR (approx 200 bp) and the poly-A tail (30 A); the dotted ellipse represents the origin of replication and the white box represents the ampicilin-resistance gene. (HPV16, human papillomavirus 16; UTR, untranslated region: bp, base pairs, A, adenine).

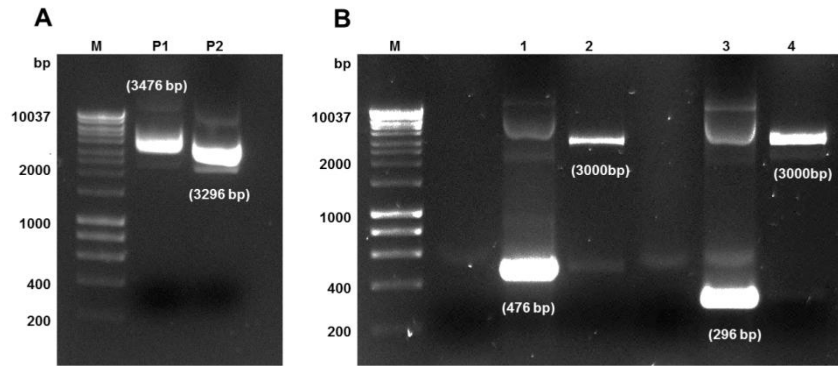

Figure 2 - Construction of the prokaryotic recombinant expression vectors of pT7TS E6 and pT7TS E7. (A) P1, pT7TS E6; P2, pT7TS E7; (B) 1, E6 PCR product; 2, enzyme digestion (BglII/EcoRV) of pT7TS E6; 3, E7 PCR product; 4, enzyme digestion (BglII/EcoRV) of pT7TS E7. (M, DNA molecular weight marker; bp, base pair).

2.3. Preparation of mRNA-based samples

mRNA encoding E6 or E7 proteins were enzymatically synthetized using MEGAscript® Kit, according to the manufacturer's recommendations. Briefly, the constructed plasmids containing E6 or E7 genes were linearized with Sma1 enzyme and concentrated to a final concentration of 1 µg/µL. Then, the linearized plasmids were *in vitro* transcribed with the reagents provided in the kit for 6h at 37ºC in order to produce the mRNA molecules. Next, the kit procedure was interrupted because it was intended to purify mRNAs by the proposed arginine affinity chromatography method. Thus, differently from the kit´s instructions, the reactions were terminated by adding 1/20th volume 0.5 M EDTA, 1/10th volume of 3M sodium acetate and two volumes of ethanol, followed by incubation at –20°C for at least 15 min. After centrifugation at 4°C for 15 min at maximum speed, the supernatant was discarded and the tubes were re-spin for a few seconds to eliminate residual fluid. Then, the pellets were solubilized in DEPC-treated water. Final mRNA-based preparations were qualitatively evaluated for purity and integrity by non-denaturing agarose gel electrophoresis and their concentration was estimated by spectrophotometric analysis. The samples were stored frozen at –80°C until be used in purification studies by arginine affinity chromatography. As a reference point, for the new purification procedure, other crude transcription reactions were concluded and purified using the conventional procedure of precipitation with LiCl. In this case, the procedure was performed according to the manufacturer's instruction of MEGAscript® Kit (Ambion). Briefly, transcription reaction was stopped and mRNAs were precipitated by adding 1/2 of transcription reaction volume of nuclease-free water and 1/2 volume of LiCl precipitation solution and mixed thoroughly. After ≥30 min incubation at –20°C, the samples were centrifuged at 4°C for 15 min at maximum speed to pellet the

mRNAs. The pellets were washed once with ~1 mL 75% ethanol and re-centrifuged. Finally, mRNA preparations were resuspended in DEPC-treated water and stored frozen at –80°C, to be used as comparison in the quality verification and concentration assessment.

2.4. Purification of mRNAs encoding E6 and E7 proteins

Chromatographic experiments were carried out using an ÄKTA Avant system with UNICORN 6 software (GE Healthcare, Uppsala, Sweden). It was used a commercial arginine–sepharose 4B gel characterized by the manufacturer as a cross-linked 4% beaded agarose matrix with a 12-atom spacer and an extent of labelling between 14 and 20 μmoL/mL. The stationary phase was packed in a 10 mm diameter×20 mm long (~2 mL) column. Temperature was maintained at 15ºC during the experiments by a circulating water bath. The system was initially prepared with ice-cold solutions of 10 mM Tris-HCl buffer, pH 8.0 in pump A and 1 M NaCl in 10 mM Tris-HCl buffer, pH 8.0 in pump B. The column was equilibrated with a mobile phase consisting of 34% of buffer B at a flow rate of 1.0 mL/min. Conductivity was rigorously controlled at 31 mS/cm. Crude transcription samples containing mRNA encoding E6 or E7 proteins were injected using a 50 µL loop at the same conditions (flow rate, mobile phase and conductivity). After 5 min run, the ionic strength of the mobile phase was increased to 100% buffer B at the same flow rate of 1 mL/min. These conditions were kept for 7 min. The chromatographic runs were monitored by an UV detector at 260 nm. To identify the eluting species, peak fractions were pooled according to the chromatograms obtained and comprising the full amplitude of the peak. After each run, the column was washed with DEPC-treated water and re-equilibrated with equilibration buffer for 20 min. For the identification of the nucleic acid species eluting in each peak, the samples were resolved on a 1.2% native agarose gel in TAE buffer (40 mM Tris base, 20 mM acetic acid and 1 mM EDTA, pH 8.0) prepared in DEPC-treated water. The gels were stained with GreenSafe (0.5 µg/mL), and photographed. The optical density of the crude transcription samples and pooled fractions was determined at 260 nm using a Nanodrop spectrophotometer to assess their quantity. The quality characterization of mRNA molecules was further performed as described below.

2.5. Protein contamination

Protein contamination in mRNA samples resulting either from purification with arginine-agarose support or purified by LiCl, was assessed using the micro-BCA (bicinchoninic acid) assay (Thermo fisher scientific Inc.), according to manufacturer's instructions. Briefly, the calibration curve was prepared using BSA standards (0.01–0.1 mg/mL). 25 µL of each standard or RNA sample were added to 200 µL of BCA reagent in a microplate and incubated for 30 min at 60 ºC. Absorbance was measured at 570 nm in a microplate reader.

2.6. Plasmid DNA template contamination

pT7TS E6 or pT7TS E7 contamination was evaluated in mRNA samples purified by arginine matrix or LiCl precipitation in an iQ5 Multicolor Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, US). Specific primers to HPV16 E6 or HPV16 E7 (Table 1) were used for the template amplification. The real-time PCR efficiency was determined for both primers using serial dilutions of mRNA samples (1:1, 1:10, 1:100 and 1:1000). Real-time PCR reactions were carried out using 1 µL of all mRNA preparations in a 20 µL reaction mix containing 10 µl Maxima SYBR Green/ROX qPCR Master mix (Fermentas, Thermo fisher scientific Inc.) and 300 nM HPV16 E6 or HPV16 E7 primers. All reactions were performed in 96-well plates for PCR heat-sealed with heat sealing film (Bio-Rad Laboratories, Hercules, CA, US). After an initial denaturation at 95ºC for 5 min, cycling conditions were used as follows: 40 cycles consisting of denaturation at 95ºC for 30 s, annealing at 60ºC for 30 s and extension at 72ºC for 1 min. Final elongation was at 72ºC for 5min. The amplified PCR fragments were checked by melting curves: reactions were heated from 55 to 95ºC with 10 s holds at each temperature (0.05ºC/s). Samples were run in triplicate for each assay. Results were analysed using iQ5 optical system software version 2.0 after manual adjustment of the baseline and fluorescence threshold.

Table 1 Primer sequences of E6 and E7 genes used for amplification. BglII and EcoRV restriction sites are shown in bold and the start and stop codons are underlined.

Primer	Sequence	Restriction site
HPV16 E6	Sense: 5'-CC**AGATCT**<u>ATG</u>CACCAAAAGAGAACTGC-3'	BglII
	Antisense: 5'-CC**GATATC**TTACAGCTGGGTTTCTCTAC-3'	EcoRV
HPV16 E7	Sense: 5'-CC**AGATCT**<u>ATG</u>CATGGAGATACACCTAC-3'	BglII
	Antisense: 5'-CC**GATATC**<u>TTA</u>TGGTTTCTGAGAACAGA-3'	EcoRV

3. Results and Discussion

The newly explored approaches of mRNA-based gene therapy are rapidly gaining consistency in cancer immunotherapy and in treating infection diseases. Furthermore, therapeutic mRNA can be easily produced and up scaled to gram quantities for clinical trials within 3 months (Pascolo, 2006), which is an advantage considering the number of patients that can be treated following this approach. While the production process of mRNA vaccines under current good manufacture practices (cGMP) is advantageously described for simplicity, the purification strategies encompass various intermediate steps of precipitation with LiCl and phenol/chloroform extractions. In addition, a final polishing step with affinity chromatography is usually used, however, the procedure is not clearly described (Pascolo, 2006). Therefore in this work, we describe a new affinity method in which mRNA molecules can be directly purified from crude transcription reactions using amino acid-based affinity chromatography with arginine as a biospecific ligand, eliminating additional steps and improving economics. Figure 3 schematically shows an overview of all the procedures employed in this study in order to demonstrate the applicability of the proposed purification method with arginine affinity chromatography.

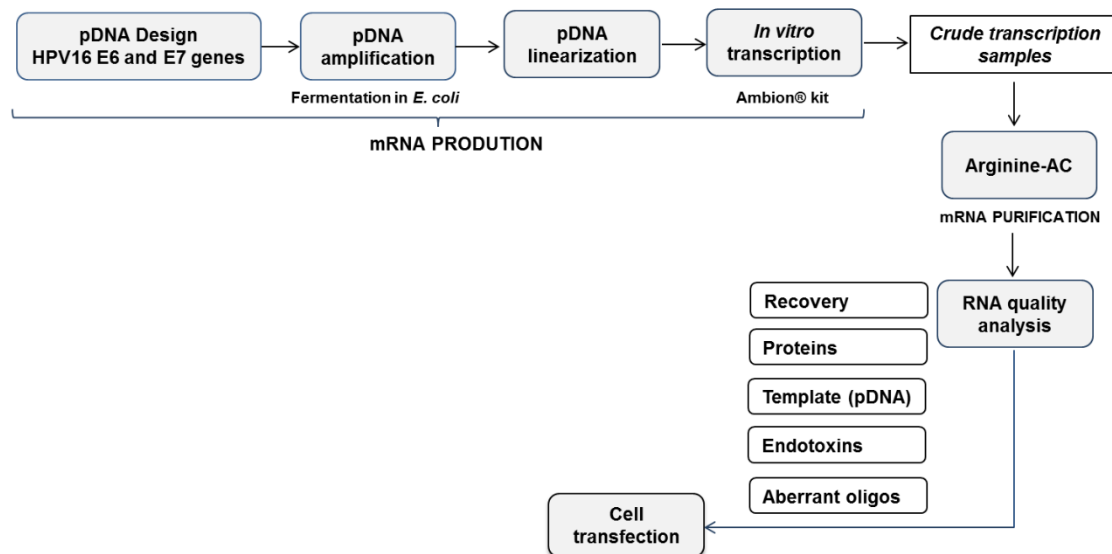

Figure 3 - General procedure of the new methodology developed for purification of E6 and E7 mRNA from crude transcription reactions. The process includes the design of a vector encoding for HPV16 E6 and E7 proteins and amplification in *E. coli*. The plasmid is extracted from the host and linearized to facilitate *in vitro* transcription of the target mRNA. Next, crude transcription samples are purified by affinity chromatography with arginine as immobilized ligand. After purification, the characterization of mRNA molecules is performed in regard to recovery yield and purity. Finally, the performance of the arginine affinity chromatography-based method should be evaluated by cell transfection experiments. (pDNA, plasmid DNA; HPV16, human papillomavirus 16; *E. coli*, *Escherichia coli*; mRNA, messenger RNA).

3.1. Construction of the prokaryotic expression vectors

To synthetically produce mRNA encoding for E6 or E7 proteins, recombinant procedures were used. Figure 1 represents the pDNA backbone, pT7TS, used for cloning and recombinant production of E6 or E7 genes. This plasmid is already described in the literature for the production of mRNA for vaccination because it contains regulatory sequences, such as UTR's, and a poly(A) tail. These confer structural modifications to mRNA molecules, which improve the transcripts' stability and enhance their translation in cells. HPV16 E6 or E7 genes were obtained from the eukaryotic expression vector p1321 HPV16 E6/E7 by PCR and were then cloned into pT7TS through the two restriction sites of BglII and EcoRV to complete the design. pT7TS also presents a linearization site that is useful for *in vitro* transcription as it improves the process. The final preparation of the recombinant pT7TS E6 and pT7TS E7 were confirmed by PCR and restriction enzyme digestion as it is shown in figure 2.

3.2. *In vitro* transcription of mRNA molecules

After amplification of the cloned vectors in *E. coli* DH5α, the plasmids were extracted from the cells and linearized using SmaI restriction site originally present in pT7TS backbone. This step is of particular importance for *in vitro* transcription of mRNA molecules since the use of circular plasmid templates will generate extremely long, heterogeneous RNA transcripts because RNA polymerases are very processive.

Transcription is a biological process of gene expression, in which a particular segment of DNA is copied into RNA by the enzyme, RNA polymerase. Both RNA and DNA are nucleic acids, which use base pairs of nucleotides as a complementary language that can be converted back and forth from DNA to RNA by the action of the correct enzymes. During transcription, a DNA sequence is read by an RNA polymerase, which produces a complementary, antiparallel RNA strand. When the gene transcribed encodes a protein, the result of transcription is mRNA, which will then be used to create that protein via the process of translation. *In vitro* transcription techniques rely on the same principle but require a purified linear engineered DNA template containing a bacteriophage promoter sequence, ribonucleotide triphosphates, a buffer system that includes dithiothreitol and magnesium ions, and an appropriate phage RNA polymerase (Beckert and Masquida, 2011).

The mRNA molecules produced by this process always need further purification since the reagents used and other products of reaction, like aberrant oligonucleotides, can still be present in the final preparations. Figure 4A shows the electrophoretic profile of the crude transcription samples obtained in this study, which include the produced E6 and E7 mRNA molecules and also residual pT7TS template and other long molecular weight impurities resultant from the *in vitro* transcription reaction. Nonetheless, these samples did not present abortive low molecular weight oligonucleotides that typically appear in the bottom of the gel.

After purification through the currently used LiCl precipitation procedure, those impurities were no longer detected by electrophoretic analysis (figure 4B). The electrophoresis was performed under non-denaturing conditions which do not reflect the correct molecular weights of E6 and E7 mRNA molecules because of the presence of secondary structures. In fact, in long transcripts as E6 mRNA, double strands structures by self-complementation can also cause artefacts in mRNAs bands, displaying two bands instead of one as it visualized in figure 4A and B.

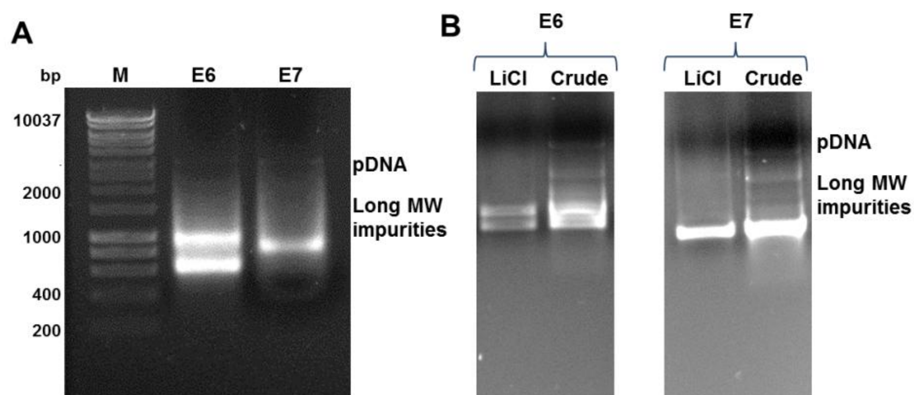

Figure 4 - Non-denaturing agarose gel electrophoresis of mRNA-based samples. (A) Samples recovered from crude transcription reactions with E6 and E7 mRNA molecules together with pDNA and other nucleotide-based impurities. (B) E6 and E7 mRNA molecules purified by LiCl precipitation compared with crude transcription samples without purification. Data are representative of several independent experiments of *in vitro* transcription reaction. (Lane M, DNA molecular weight marker; bp, base pair; pDNA, plasmid DNA template; MW, molecular weight).

3.3. Purification of mRNA molecules by arginine affinity chromatography

The crude transcription samples previously analysed were purified by using affinity chromatography with an arginine-agarose support in order to eliminate the impurities derived from the production process. Chromatographic conditions were optimized concerning mRNA retention in the column and the molecular stability. The temperature in the column was set at 15ºC and ice cold solutions of 10mM Tris-HCl, pH 8 in pump A and 1M NaCl in 10mM Tris-HCl, pH 8 in pump B were used. Different salt concentrations were tested to achieve mRNA retention, in order to exploit favoured interactions between mRNA molecules and arginine-agarose matrix that allow their purification. During these experiments it was verified the

need for strict control of chromatographic conditions in order to maintain the reproducibility, since a slight variation in conductivity (salt concentration and/or temperature) affected mRNAs retention (results not shown).

Figure 5 shows the chromatograms with the optimized elution gradient applied to arginine-agarose chromatography for the purification of E6 and E7 mRNAs and the electrophoretic analysis of the pure factions (figure 5A and B, respectively). The agarose gel electrophoresis was used to detect and identify the nucleic acids eluting in each peak. Impure E6 and E7 samples (figure 5A and 5B, lanes C) were also run on the gel for comparative purposes.

mRNAs purification by arginine-agarose support was performed in two steps. After column equilibration with 34% Pump B, crude transcription samples were injected onto the column and in the first gradient step with the same ionic strength, pT7TS template was eluted in the first peak, as it can be seen in the lane 1 of electrophoretic analyses (figure 5A and 5B). In a second step, the ionic strength was increased to 100% Pump B allowing pure mRNAs to elute, as it is observed in lane 2 (figure 5A and 5B). Purification was achieved in a single step in approximately 12 min, which greatly improves the time expended in mRNA purification when employing the multi-step procedure of LiCl.

Interestingly, the first peak, where the elution of pDNA was hardly detected by electrophoresis, due to low initial concentration in the injected transcription reactions, showed a higher peak area than the peak where mRNA molecules were eluting with much higher concentration. This difference can be explained by the fact that the electrophoresis only detects eluting nucleic acids, but other impurities from transcription process that are not detected by this electrophoresis can also absorb at 260nm. In fact, while assessing proteins contamination in mRNA preparations purified by arginine affinity chromatography, we verified that the first recovered peak contained a higher concentration of proteins than the second peak (data not shown).

The presence of two main peaks in both chromatograms indicates that nucleic acids present in the samples interact differently with arginine support. Curiously, E6 and E7 mRNAs were purified from impurities of the synthesis by using the same chromatographic conditions, which suggests that the major interactions occurring between mRNAs and the support are the same.

In affinity chromatography, the interactions occurring between a target biomolecule and its specific ligand are responsible for the high selectivity achieved in this technique. Those interactions are similar to the contacts described in many biological systems (Roque and Lowe, 2008). Thus, the binding mechanism is suggested to involve a combination of phenomenological interactions, such as biorecognition, between the amino acids and RNA, including electrostatic or hydrophobic interactions, van der Waals forces, or hydrogen bonding (Yarus *et al.*, 2009). However, depending on the environmental conditions

established, some interactions can be more favored than others. Since RNA is negatively charged due to the phosphate groups in its backbone, it is reasonable to predict a favored electrostatic interaction between mRNAs phosphate groups and arginine ligands that are positively charged at the pH in study. In fact, some molecular recognition studies reported arginine as the preferential amino acid to contact with RNA when the overall negative charge of RNA is considered (Treger and Westhof, 2001, Yarus *et al.*, 2009). Additionally, saturation transfer difference-nuclear magnetic resonance (STD-NMR) spectroscopy and surface plasmon resonance (SPR) biosensor techniques recently reported that adenine, cytosine, and guanine polynucleotides bind to arginine-agarose support mainly through the sugar-phosphate backbone (Cruz *et al.*, 2011).

On the other hand, pT7TS template is also a negative charged nucleic acid and it was observed to be less retained, revealing a not so strong affinity for the arginine-agarose support as in case of RNA transcripts. Therefore, nucleotide bases exposure is suggested as having a crucial role in nucleic acid retention. The double-stranded structure of pDNA causes the coverage of the nucleotides bases disabling the interactions with arginine-agarose matrix, while single-stranded RNA can expose them, strengthening interactions with the support. In fact, previous studies in regard to the purification of pDNA by arginine chromatography also reported similar findings (Sousa *et al.*, 2008, Sousa *et al.*, 2009).

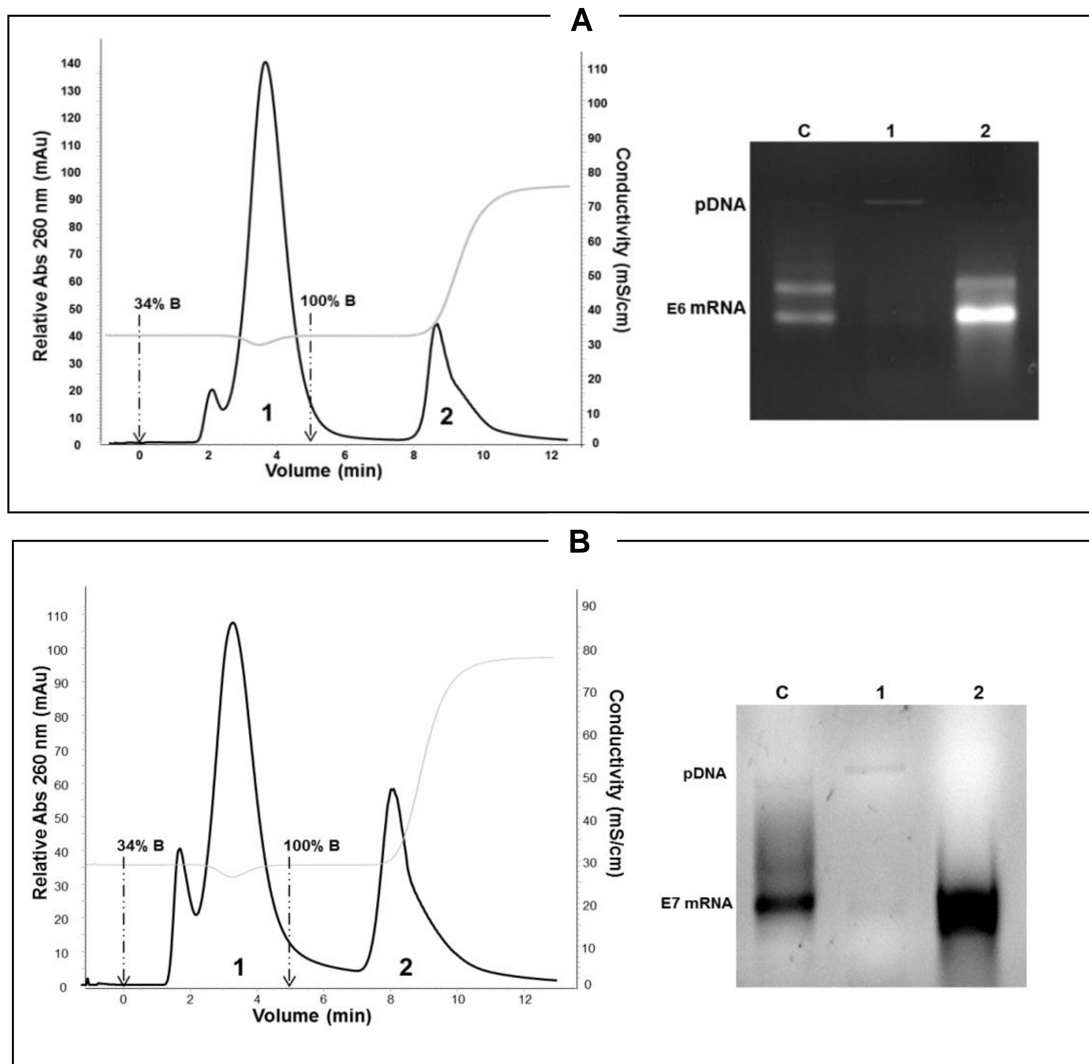

Figure 5 Arginine affinity chromatography purification of E6 (A) and E7 (B) mRNA molecules from crude transcription samples with the corresponding non-denaturing agarose gel electrophoresis of the samples collected at the column outlet. Chromatographic conditions for the mRNA purification included arginine column at 15ºC and ice cold solutions. Solution in pump A was 10mM Tris-HCl, pH 8 and in pump B 1M NaCl in 10mM Tris-HCl, pH 8. Elution was performed at 1.0 mL/min by stepwise increasing NaCl concentration in the eluent from 34% pump B to 100% pump B, as represented in the arrows. The conductivity was followed along the chromatographic purification as indicated by the light grey line. Fractions corresponding to peaks (1) and (2) are shown in lanes 1 and 2, respectively. Lane C, crude transcription samples injected onto the column. Data are representative of several independent experiments from different *in vitro* transcription reaction. (Abs, absorbance; nm, nanometers, pDNA, plasmid DNA template; mRNA, messenger RNA).

3.4. mRNA quality characterization

To certify the efficacy and applicability of arginine affinity method in the purification of mRNA molecules, the recovery yield and the quality of the preparations in regard to purity, were evaluated. Recovery yields were semi-quantitatively assessed by spectrophotometric analysis, measuring absorbance at 260 nm of the resulting peaks. Moreover, the quality of mRNA preparations was evaluated concerning total protein content, determined by micro-BCA assay, pDNA template contamination identified by real-time PCR and aberrant oligonucleotides detected by electrophoretic analysis. Crude transcription reactions purified by LiCl precipitation were used as a reference point for the new purification procedure in order to control and compare the quality of mRNA prepared by both methods.

Table 2 summarizes the recovery yields and total protein content. Spectrophotometric quantification demonstrated a high recovery yield for both E6 and E7 mRNAs purified by the arginine-based process, achieving 97.5% and 98% of recovery, respectively. Protein contamination was also determined and compared with samples precipitated with LiCl. The results indicated that arginine affinity chromatography efficiently purified mRNA with low protein content. However, while E6 mRNA purified by arginine-agarose matrix presented lower protein levels than precipitated with LiCl, E7 mRNA had similar protein concentration in both arginine-based or control preparations.

pT7TS template contamination was also assessed in mRNA samples purified by arginine-based method or precipitated by LiCl. Figure 6 shows the good efficiencies (around 100%) of the primers used to amplify pDNA. The results of real-time PCR were based on the evaluation of threshold cycles (Ct), which is defined as number of amplifications cycles necessary to detect molecules on a sample (Schmittgen and Livak, 2008). Thus, the results showed no detection of pDNA in samples purified by arginine affinity chromatography however, pDNA was detected in samples precipitated with LiCl after 22 cycles of amplification.

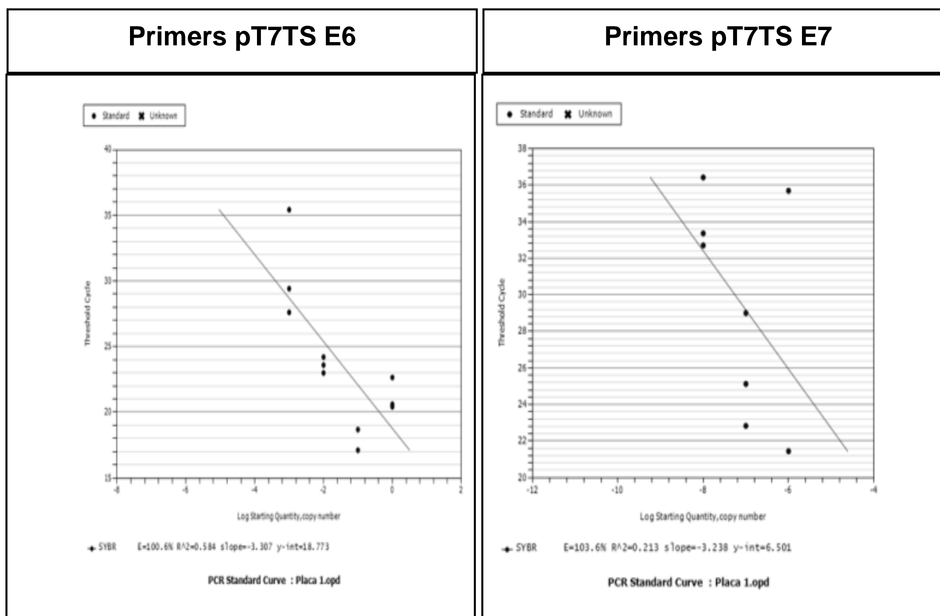

Figure 6 - Real time PCR output showing PCR efficiencies of the primers used to detect pT7TS E6 and E7 templates in mRNA preparations. (mRNA, messenger RNA).

During all the methodology employed in this study, electrophoretic analyses were used as complementary technique in order to detect and identify the nucleic acids in each sample (figures 4 and 5), but it also enabled the verification of aberrant oligonucleotides, since they can be detected by the presence of bands with higher or lower molecular weight than E6 or E7 mRNAs (Beckert and Masquida, 2011). While electrophoresis performed in crude transcription reactions demonstrated the presence of high molecular weight oligonucleotides impurities (figure 4A and B, lane Crude), the purification by arginine affinity chromatography or LiCl eliminated the presence of aberrant oligonucleotides (figure 4B, lane LiCl and 5A and B, lane 2). Additionally, this technique can also provide important qualitative information on RNA integrity. Typically, electrophoretic profile of degraded RNA presents a smear-like band instead of a compact band. Our results demonstrate an accurate banding profile for mRNA molecules, suggesting that mRNAs are chemically intact. Although a resultant smear can be occasionally visualized together with mRNA bands (figure 5, lane 2), that effect is not due to sample degradation but instead it reflects persistent mRNA secondary structure, since electrophoretic analyses are performed under native conditions (Farrell, 2005).

In order to render arginine affinity chromatography as an effective methodology for mRNA purification pursuing vaccine-based therapies, endotoxins content together with cell transfection experiments should also be considered essential parameters.

Table 2 - Characterization of mRNA preparations in relation to recovery yield and protein contamination. Protein concentration is compared in samples purified by the new arginine affinity chromatography method or the conventional LiCl purification. The correlation coefficient of protein calibration curve was 0.998. Recovery yield data is presented as means with SD (n=3). (mRNA, messenger RNA; Arg-AC, arginine affinity chromatography; LiCl pp, lithium chloride precipitation; ND, not detected by gel electrophoresis analysis).

Method	Sample		Volume (µL)	Impurities (µg)	mRNA (µg)	Recovery yield (%)	Proteins (µg/µL)
Arg-AC	E6	Crude samples	1000	200.0±21		97.5±17	26.2
		Peak 1	500	72.4±29	ND		3.1
		Peak 2	500	ND	122.6±11		1.6
	E7	Crude samples	1000	200±14		98±14	27.6
		Peak 1	500	62.0±19	ND		11.6
		Peak 2	500	ND	134.0±17		7.6
LiCl pp	E6		200	30.5±25		____	13.2
	E7		200	22.7±18		____	8.0

4. Conclusion

The present study introduces a new affinity chromatography method using arginine as a biospecific ligand for mRNA purification, which reveals several advantages over the existing procedures, namely LiCl precipitation. The exploitation of specific affinity interactions between RNA and arginine allowed obtaining mRNA preparations with high recovery yield and good purity, eliminating additional steps and improving global economics of the production process. The results obtained so far indicated that the chromatographic separation is a promising strategy for mRNA purification pursuing therapeutic applications.

Acknowledgments

The authors acknowledge to Prof. Peter Howley from Harvard Medical School, Boston, USA and to Prof. Paul Krieg from University of Arizona College of Medicine, USA for kindly provide their plasmids p1321 HPV-16 E6/E7 and pT7TS, respectively. This work was supported by FCT, the Portuguese Foundation for Science and Technology (PTDC/EBB-BIO/114320/2009; PEst-C/SAU/UI0709/2011 COMPETE and EXPL/BBB-BIO/1056/2012, COMPETE: FCOMP-01-0124-FEDER-027560). Rita Martins also acknowledges a fellowship (SFRH/BD/ 64100/2009) from FCT.

The authors declared no conflict of interest.

References

Beckert B, Masquida B. Synthesis of RNA by in vitro transcription. Methods Mol Biol. 2011; 703:29-41.

Bharti AC, Shukla S, Mahata S, Hedau S, Das BC. Anti-human papillomavirus therapeutics: facts & future. Indian J Med Res. 2009; 130:296-310.

Bringmann A, Held SA, Heine A, Brossart P. RNA vaccines in cancer treatment. J Biomed Biotechnol. 2010; 2010:623687.

Cruz C, Cabrita EJ, Queiroz JA. Screening nucleotide binding to amino acid-coated supports by surface plasmon resonance and nuclear magnetic resonance. Anal Bioanal Chem. 2011; 401:983-93.

Farrell RE. RNA Methodologies: A Laboratory Guide for Isolation and Characterization. 3rd ed. New York: Elsevier Academic Press; 2005.

Franco EL, Coutlee F, Ferenczy A. Integrating human papillomavirus vaccination in cervical cancer control programmes. Public Health Genomics. 2009; 12:352-61.

Huh WK, Alvarez, R. D., Curiel, D. T. Gene therapy and cervical cancer: an overview. CME Journal of gynecologic oncology. 2001; 6:377-82.

Kreiter S, Diken M, Selmi A, Tureci O, Sahin U. Tumor vaccination using messenger RNA: prospects of a future therapy. Curr Opin Immunol. 2011; 23:399-406.

Martins R, Maia CJ, Queiroz JA, Sousa F. A new strategy for RNA isolation from eukaryotic cells using arginine affinity chromatography. J Sep Sci. 2012a; 35:3217-26.

Martins R, Queiroz JA, Sousa F. A new affinity approach to isolate *Escherichia coli* 6S RNA with histidine-chromatography. J Mol Recognit. 2010; 23:519-24.

Martins R, Queiroz JA, Sousa F. Histidine affinity chromatography-based methodology for the simultaneous isolation of *Escherichia coli* small and ribosomal RNA. Biomed Chromatogr. 2012b; 26:781-8.

Martins R, Queiroz JA, Sousa F. New approach in RNA quantification using arginine-affinity chromatography: potential application in eukaryotic and chemically synthesized RNA. Anal Bioanal Chem. 2013; 405:8849-58.

Munger K, Phelps WC, Bubb V, Howley PM, Schlegel R. The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. J Virol. 1989; 63:4417-21.

Pascolo S. Vaccination with messenger RNA. Methods Mol Med. 2006; 127:23-40.

Pascolo S. Vaccination with Messenger RNA (mRNA). In: Bauer S, Hartmann G, editors. Toll-Like Receptors (TLRs) and Innate Immunity: Springer Berlin Heidelberg; 2008. p. 221-35.

Ponsaerts P, Van Tendeloo VF, Berneman ZN. Cancer immunotherapy using RNA-loaded dendritic cells. Clin Exp Immunol. 2003; 134:378-84.

Roque AC, Lowe CR. Affinity chromatography: history, perspectives, limitations and prospects. Methods Mol Biol. 2008; 421:1-21.

Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. Nat Protoc. 2008; 3:1101-8.

Sousa F, Matos T, Prazeres DM, Queiroz JA. Specific recognition of supercoiled plasmid DNA in arginine affinity chromatography. Anal Biochem. 2008; 374:432-4.

Sousa F, Prazeres DMF, Queiroz JA. Improvement of transfection efficiency by using supercoiled plasmid DNA purified with arginine affinity chromatography. J Gene Med. 2009; 11:79-88.

Thornburg C, Boczkowski D, Gilboa E, Nair SK. Induction of cytotoxic T lymphocytes with dendritic cells transfected with human papillomavirus E6 and E7 RNA: implications for cervical cancer immunotherapy. J Immunother. 2000; 23:412-8.

Treger M, Westhof E. Statistical analysis of atomic contacts at RNA-protein interfaces. J Mol Recognit. 2001; 14:199-214.

Van Lint S, Heirman C, Thielemans K, Breckpot K. mRNA: From a chemical blueprint for protein production to an off-the-shelf therapeutic. Hum Vaccin Immunother. 2013; 9:265-74.

Vergati M, Intrivici C, Huen N-Y, Schlom J, Tsang KY. Strategies for Cancer Vaccine Development. J Biomed Biotechnol. 2010; 2010.

Weide B, Garbe C, Rammensee HG, Pascolo S. Plasmid DNA- and messenger RNA-based anti-cancer vaccination. Immunol Lett. 2008; 115:33-42.

Yarus M, Widmann JJ, Knight R. RNA-amino acid binding: a stereochemical era for the genetic code. J Mol Evol. 2009; 69:406-29.

Chapter 4

Concluding remarks

Although over the recent years therapeutic potential of RNA has been evidenced, intensifying the studies on RNA in many research areas, various limitations have been found regarding the use of suitable techniques to isolate and purify RNA molecules.
Due to the particular chemical structure of RNA, its formulations are prepared with methods that employ chemical solvents and denaturing conditions to ensure chemical stability and enhanced purification. However, these procedures turned out to be inadequate to recover a product to be biologically applied and do not conduct to an efficient, tolerable and scalable process.

This thesis established a general hypothesis of using amino acid-based affinity chromatography to specifically and accurately purify RNA molecules by exploiting the interaction between the histidine or arginine amino acids and RNA, in order to overcome the challenges in RNA purification and to improve the purification process. The progress of this work disclosed several interesting characteristics of RNA molecules, including chromatographic behavior and natural interactions that can occur between amino acids-based supports and RNA molecules, which evidenced the remarkable versatility of amino acids-based affinity chromatography in the purification and analysis of RNA.

Histidine-agarose was the first support to be tested in RNA purification by exploiting the interaction that might accrue with RNA molecules from *E. coli*. Because of the presence of the aromatic imidazole ring in histidine structure the interactions between RNA and the matrix were suggested to be mainly hydrophobic either by ring-staking or hydrogen bonds, thus ammonium sulphate gradients were used in histidine strategies.

As described in paper I, the first studies employed the sRNA class because of their increasing importance as regulatory molecules in gene expression in all organisms and thus, the availability of techniques to accurately purify these molecules would be of great benefit for their research. Because of the several limitations in the extraction procedures to obtain the sRNA samples to be used in the chromatographic studies, namely the deficiency of convenient techniques to fractionate RNA classes or a single RNA type, a first scheme to obtain an isolated pool of sRNA had to be developed based on a precipitation with ammonium sulphate. Consequently, the application of histidine affinity chromatography showed a specific recognition between the histidine support and 6S RNA, which allowed its selective isolation from the complex mixture of other sRNAs, demonstrating the potential of this technique in RNA purification purposes.

Driven from the results and conclusions obtained in the first study, we found necessary to explore the use of histidine affinity chromatography in a method allowing the isolation and purification of the main RNA classes from the impurities of host cells without the use of hazardous extractions (paper II). A lysis step was successfully optimized to extract the nucleic acids from the cells, avoiding the use of phenol/chloroform, and a preliminary separation between the sRNA and rRNA classes in the cell lysates was positively accomplished by the precipitation with ammonium sulphate (developed in the previous work). But, the performance of histidine matrix was highlighted, since it had the ability to specifically interact with sRNA or rRNA, eliminating genomic DNA and proteins and preserving the integrity of RNA molecules. Therefore, this method allowed the preparation of both RNA classes with high purity degree and recovery yields.

The versatility of histidine matrix in both strategies suggested that the underlying mechanism in the purification of single 6S RNA molecule or both RNA classes involves not only hydrophobic interactions, but also a biorecognition of RNA bases by histidine. Therefore, histidine affinity chromatography represents a major advance in the ability to accurately purify RNA molecules, since the conventional procedures do not have the capacity to fractionate RNA at a single molecule and the simultaneous isolation of sRNA and rRNA in a single procedure is not available yet.

Despite the effectiveness of histidine-based methodologies, the requirement for high salt concentration can be a disadvantage, especially with regard to biotechnological application, because the use of salt is associated with higher costs and environmental impact. The use of affinity supports with arginine as the immobilized ligand could overcome this limitation and improve RNA purification. Since arginine is a positively charged amino acid, electrostatic interactions could be exploited using only mild condition with biological salts as sodium chloride. Moreover, arginine-base interactions have been recognized as the most prevalent interactions in several protein-RNA complexes.

In paper III, arginine affinity chromatography was used to develop a strategy involving RNAt purification from impurities of eukaryotic cell extracts to aid in molecular biology analysis, such as gene expression, since these experiments employ high amounts of RNAt and are critically dependent on the quality of RNA preparations. In this study, arginine support demonstrated an exceptional ability to interact with all functional classes of RNA (RNAt), despite their structural diversity and different folding states as they are in their native state. These strong and selective interactions with the affinity matrix were suggested to be due to the multiplicity of interactions that arginine side chain is able to establish, promoting, in this case, a multi contact with RNA backbone or RNA bases, according to RNA folding. Those properties enabled a high recovery yield, a high integrity and good purity of RNA preparations. The successful applicability of arginine-based technique in the purification of

RNAt attested by gene expression analysis, proved the reliability and potential of arginine support in the purification of RNA.

Taken together the results of paper I, II and III, the RNA isolation and purification from a complex biological mixture using the amino acids-based affinity techniques showed several improvements over the currently phenol/chloroform or SPE isolation methods. Moreover, this method simplified the workflow integration and miniaturizes sample handling process, making them useful for the development of a RNase and organic solvents free methodology. The developed approaches can be very supportive in structural and functional studies and clinical investigations.

Furthermore, due to the multiple interactions that arginine matrix evidenced in the previous work enabling the accurate purification of RNAt, an analytical technique for quantification and quality assessment of RNAt was also developed and validated, according to international and European legislation for bioanalytical methods (paper IV). The development of this approach with arginine support was imperative, since several flaws still exist in the current RNA quantification techniques, as the lack of selectivity. The versatility of arginine support was demonstrated by its applicability in the quantification of RNA from different eukaryotic cells and in crude samples of synthetized RNA, which highlighted the potential multipurpose applicability in molecular biology RNA-based analysis and RNA therapeutics.

The reliability and potential of arginine-based chromatography in RNA purification triggered the development of a new approach in order to render this technique useful in therapeutic application of RNA (paper V). This work showed the singular ability of arginine-based chromatography in purifying mRNA molecules from synthetic crudes of *in vitro* transcription reactions, pursuing mRNA vaccination for cervical cancer. Moreover, it advantageously eliminated additional steps and improved the global economics of the production process, proving to be an interesting option for the downstream processing of RNA therapeutic products.

Overall, this thesis shows that the hypothesis of using affinity chromatography based on natural occurring biological interactions between amino acids and RNA, to develop new improved approaches for RNA isolation and purification was successfully accomplished. This work provided new insights in the way RNA can be accurately purified, expectantly contributing for the future development of new and more robust bioseparation methods.

Future perspectives

Despite all the efforts performed during this thesis, further investigations is still need to increase the significance of these important findings and to realize the development of new and more robust purification methods for RNA.

As future perspectives and following the work already performed, we propose more studies concerning the improvement on affinity ligands and the capacity of supports. The progression on synthetic synthesis and chemical modification would promisingly aid in the development of highly selective and stable synthetic affinity ligands. Furthermore, ready available technologies as monoliths could be experienced as affinity chromatographic supports based on the immobilization of amino acids, which would potentially improve the speed, resolution and capacity of the chromatographic operation while combining the distinctive property to simulate and exploit natural biological interactions to selectively purify RNA. This should largely benefit the industrial production of RNA biotherapeutics.

Moreover, complementary studies on the understanding of the biorecognition phenomenon between the amino acids and RNA molecules would be of great value, since the knowledge of the involved interactions in the specific recognition of RNA and the understanding of the chromatographic conditions responsible by favoring or not the underling interactions could more easily permit the control and manipulation of the selective retention and/or elution of the different RNA molecules and impurities. Accordingly, the molecular identification of those interactions could be improved through the determination of affinity constants, and characterization of the nucleotide regions involved in these interactions by SPR-biosensor and NMR spectroscopy.

Finally, the applicability of this work would greatly increase if RNA biotherapeutic product was tested in *in vivo* systems, becoming interesting to further evaluate the therapeutic function.