ROMA TRE
UNIVERSITÀ DEGLI STUDI

DOTTORATO DI RICERCA IN BIOLOGIA

Scienze Biomolecolari e Cellulari

XXIII CICLO

Investigation on the biogenesis of macromolecules in prebiotic environments

Studio sulla biogenesi di macromolecole in ambiente prebiotico

Rafał Maciej Wieczorek

A.A. 2010/2011

Docente Guida: Prof. Pier Luigi Luisi

Coordinatore: Prof. Paolo Mariottini

Acknowledgments

This work would not be possible without the help I received from so many people, whether in Rome, in Venice or in Odense.

Prof. Pier Luigi Luisi was my guide to the field of the origin-of-life as well as the author of the exciting and challenging idea behind this work. Prof. Pierre-Alain Monnard provided solution when the project was facing a dead end. In Venice Davide De Lucrezia helped me with many issues in the beginning of my doctorate. Katarzyna Adamala from Boston was discussing with me all SerHis related matters throughout this three years. Pasquale Stano from Rome had numerous insightful comments regarding my work and most important was the person that kept all my doctoral paper work at check when I was away from Rome for long periods. In Odense Phillip Löffler was the person that introduced me to the eutectic RNA systems. Another Odense chemist Mark Dörr was always eager to discuss and comment on the progress of my experiments.

I have also recieved help and inspiration from many other members of RomaTre group, ECLT group and FLinT group.

My wife, Rachel, was always supporting me throughout all these years.

Riassunto

La chimica prebiotica è un campo nel quale si cerca di ricreare il percorso di sintesi chimica che ha portato alla formazione di molecole biologiche, e in ultima analisi all'origine della vita. Negli esperimenti di chimica prebiotica (utilizzando condizioni che erano presenti sulla Terra prima della comparsa della vita primitiva) i ricercatori sono stati in grado di sintetizzare amminoacidi e peptidi brevi, ma non acidi polinucleici e mononucleotidi.

L'attività catalitica di diversi peptidi brevi è stata recentemente descritta in letteratura. Alcuni di questi peptidi sono sufficientemente brevi da essere considerati prebioticamente plausibili. Uno di questi è il dipeptide SerHis (Li *et al.* 2000), che ha un'ampia attività idrolitica ed e' in grado di idrolizzare proteine, acidi nucleici ed esteri. In virtù delle sua attività catalitica nelle reazioni di idrolisi, il dipeptide SerHis deve necessariamente poter catalizzare la reazione inversa, cioè la formazione di legami esterei, fosfodiesterei e peptidici. Infatti, è ben noto che le reazioni di condensazione e quelle di idrolisi sono l'una l'opposto dell'altra, e possono essere catalizzate da una stessa specie chimica. Ciò è già stato dimostrato nel caso di enzimi come la chimotripsina. Pertanto, ci si aspetta che SerHis sia in grado di catalizzare la condensazione tra opportuni reagenti, laddove le condizioni termodinamiche diventino favorevoli alla formazione di un prodotto di condensazione. Questa ipotesi è stata verificata sperimentalmente in un lavoro svolto nel gruppo di ricerca in cui ho svolto la tesi, dove è stato provato che SerHis catalizza la formazione di legami peptidici tra amminoacidi attivati (Gorlero *et al.* 2009).

Lo scopo di questo lavoro è stato quello di esplorare le capacita' di catalisi da parte del dipeptide SerHis nei confronti della oligomerizzazione di acidi nucleici, principalmente RNA. Per indagare le condizioni richieste per questa reazione, sono state studiate la reazione di condensazione, le proprietà del catalizzatore, e il possibile meccanismo di catalisi.

L'ambiente utilizzato per le reazioni qui descritte consiste in un eutettico. La miscela eutettica utilizzata in questo lavoro è un sistema bifasico composto da una soluzione acquosa dei reagenti e da ghiaccio, mantenuti alla temperatura di -18 °C. Dal punto di vista fisico, una miscela eutettico è composta da microcanali di soluzione liquida dispersi tra estese zone di ghiaccio. In tali condizioni, compatibili con possibili scenari della

Terra primordiale, si ottiene la compartimentalizzazione dei soluti nella microfase liquida con conseguente sovraconcentrazione dei soluti, che vengono esclusi dal ghiaccio. Inoltre, a causa del ridotta attività termodinamica dell'acqua nella fase liquida, tale microcompartimentalizzazione favorisce reazioni di condensazione (che implicano una perdita di acqua come prodotto di reazione) come quelle richieste per la formazione di acidi nucleici promossa dal dipeptide SerHis.

La catalisi della reazione di polimerizzazione degli acidi nucleici ha avuto esiti positivi. In condizioni standard, la resa massima di polimerizzazione ottenuta con il catalizzatore SerHis era intorno al 10%. La reazione decorre con successo in un ampio range di parametri sperimentali, che riflettono diverse condizioni ambientali plausibili prebioticamente: in diverse condizioni di pH (5,0-8,2), in tamponi diversi, con diversi derivati dell'imidazolo, e utilizzando diversi monomeri di RNA e DNA.

Le proprietà di SerHis dipendono dalla disposizione strutturale dei suoi gruppi funzionali; sorprendentemente il dipeptide isomerico HisSer non ha un'attività catalitica misurabile. L'attività di SerHis è rilevabile fino a concentrazioni molto basse - 8 μM, mille volte inferiore alla concentrazione del substrato. Inoltre, è stato dimostrato che SerHis è capace di consentire un meccanismo a turnover tipico degli enzimi.

Per quanto riguarda lo studio del meccanismo, è stato isolato un prodotto intermedio e caratterizzato attraverso la spettrometria di massa. Risulta essere un addotto covalente tra SerHis e il substrato di reazione (un mononucleotide attivato). Questo intermedio, posto nuovamente a reagire, si decompone in SerHis (che continua a catalizzare la reazione), e oligomeri di acidi nucleici. Sulla base dei dati ottenuti, sono state fatte delle ipotesi sulla struttura chimica dell'intermedio e sul meccanismo di reazione.

Il fatto che SerHis può catalizzare la formazione di RNA è di grande importanza nel campo della chimica prebiotica. Infatti, un notevole ostacolo all'ipotesi del mondo RNA deriva dalla non plausibilità delle vie sintetiche di oligo- e polimerizzazione dell'RNA. Pertanto il suo emergere resta una domanda aperta nel campo dell'origine della vita. Come mostrato in questa tesi, la capacità di piccoli peptidi (che sono in effetti prodotti prebiotici) di oligomerizzare o polimerizzare l'RNA permette di proporre uno scenario verosimile che collega la chimica prebiotica e il mondo RNA. Si può quindi immaginare, in linea di principio, un percorso graduale dalla chimica prebiotica della sintesi spontanea di amminoacidi e di piccoli

peptidi, al mondo dell'RNA, e quindi verso la vita cosi come lo conosciamo oggi. E' importante notare che così come piccoli peptidi possono condensare nucleotidi a dare oligonucleotidi, è plausibile pensare che più frammenti di oligonucleotidi possano combinarsi a dare polinucleotidi, in modo da ottenere lunghe catene di acido ribonucleico.

Inoltre, questo studio ci consente una maggior compresione del fenomeno dell'origine della vita e delle origini della catalisi enzimatica. Questo lavoroha provatoche un peptide di due soli amminoacidi puo' già presentare proprietà enzimatiche. Ciò dimostra che la selezione verso gli enzimi - catalizzatori altamente efficiente, potrebbe aver avuto inizio in una fase molto precoce di evoluzione chimica.

Summary

Prebiotic chemistry is a field which attempts to recreate the pathways of chemical synthesis which led to the formation of biological molecules, and ultimately the origin of life. In experiments of prebiotic chemistry (using conditions that existed on Earth before the appearance of primitive life), researchers were able to synthesize amino acids and short peptides, but never nucleic acids and mononucleotides.

The catalytic activity of several short peptides was recently described in literature; some of these peptides are short enough to be considered prebiotically plausible. One of these is the dipeptide SerHis (Li *et al.* 2000), which has broad hydrolytic activity and is capable of hydrolyzing proteins, nucleic acids and esters. By virtue of its catalytic activity in hydrolysis reactions, the dipeptide SerHis must necessarily be capable of catalyzing the reverse reaction, *i.e.* the formation of ester-linkages, phosphodiesters and peptides. In fact, it is well known that the condensation reactions and the hydrolysis are each the opposite of the other, and can be catalyzed by a single chemical species. This has already been demonstrated in the case of enzymes such as chymotrypsin.

Therefore, it is expected that SerHis is capable of catalyzing the condensation between the appropriate reagents, where the thermodynamic conditions are favorable for the formation of a condensation product. This hypothesis was verified experimentally in work done in the research group where I did my thesis, where it was proven that SerHis catalyzes the formation of peptide bonds between activated amino acids (Gorlero *et al.* 2009).

The purpose of this study was to explore the capabilities of catalysis by the dipeptide SerHis of the oligomerization of nucleic acids, primarily RNA. To investigate the conditions required for this reaction, we have studied the condensation reaction, the properties of the catalyst, and the possible mechanism of catalysis.

The environment used for the reactions described here consists of an eutectic phase. The eutectic mixture used in this study is a two-phase system consisting of an aqueous solution of reagents and ice, kept at -18 °C. From a physical point of view, an eutectic mixture is composed of micro-dispersed liquid solution including extensive areas of ice. Under these

conditions, consistent with scenarios of the early Earth, the compartmentalization of solutes in the liquid results in the upconcentration of solutes, which are excluded from the ice. Moreover, because of the reduced thermodynamic activity of water in liquid phase, this microcompartimentalization promotes condensation reactions (involving a loss of water as a reaction product) as those required for the formation of nucleic acids promoted by the SerHis dipeptide.

The catalysis of the polymerization reaction of nucleic acids was successful; in standard conditions, the maximum yield obtained by the polymerization catalyst, SerHis, was around 10%. The reaction runs successfully in a wide range of experimental parameters, which reflect several plausible prebiotic environmental conditions: in different conditions of pH (5.0 to 8.2), in different buffers, with different imidazoles, and using various monomers of RNA and DNA.

The catalytic properties of SerHis depend on the structural arrangement of its functional groups, strikingly, the isomeric dipeptide HisSer has no measurable catalytic activity. SerHis activity is detectable down to very low concentrations - 8 μM - a thousand times smaller than the substrate concentration. In addition, it was shown that SerHis is capable of multiple turnovers - a mechanism typical of enzymes.

Regarding the study of the mechanism, an intermediate product was isolated and characterized by mass spectrometry. It appears to be a covalent bond between SerHis and the substrate from the reaction (activated mononucleotide). This intermediate, when allowed to react again, decomposes into SerHis (which continues to catalyze the reaction), and nucleic acid oligomers. On the basis of the data obtained, hypothesis for the chemical structure of the intermediate and the mechanism of the reaction were formed.

The fact that SerHis can catalyze the formation of RNA is of great importance to the field of prebiotic chemistry. In fact, a major obstacle in the RNA world hypothesis stems from the implausibility of synthetic pathways of oligo- and poly-merization of RNA. Accordingly, its emergence remains an open question in the origin of life. As shown in this thesis, the ability of small peptides (which are prebiotic) to perform the oligomerization or polymerization of RNA can offer a plausible scenario that links the chemical and prebiotic RNA world. You can then imagine, in principle, a step by step pathway from the spontaneous chemical prebiotic

synthesis of amino acids and small peptides to the RNA world, and then to life as we know it today. It is important to note that just as small peptides can catalyze the formation of oligonucleotides from nucleotides, it is conceivable that several fragments of oligonucleotides can then be combined to give polynucleotides, in order to obtain long chains of ribonucleic acid.

In addition, this study allows for a greater understanding of the origin of life and the origins of enzyme catalysis. This work has shown that even a peptide of only two amino acids can present enzymatic properties. This shows that the selection which led to enzymes - highly efficient catalysts - may have begun at a very early stage of chemical evolution.

List of abbreviations

All discussed amino acids are in the L configuration.

2-MeImpA – adenosine-5'-(2-methylimidazole) monophosphate
2-MeImpG – guanosine-5'-(2-methylimidazole) monophosphate
2-NH_2-BenzImpG – guanosine-5'-(2-amino benzimidazole) monophosphate
2,4-diMeImpG – guanosine-5'-(2,4-dimethylimidazole) monophosphate
A – adenine
Ac – acetyl group
Ala – alanine
Asp – asparagine acid
ATP – adenosinotriphosphate
C – cytosine
COS – carbonyl sulfate
Cys – cysteine
DMF – dimethylformamide
DMSO – dimethylsulfoxide
DNA – deoxyrybonucleic acid
Et – etyl group
Et_3N – N,N,N-triethylamine
Et_2O – diethyl ether
G – guanine
Glu – glutamic acid
Gly – glicyne
GMP – guanine monophosphate
GppG – guanosine pyrophosphate dimer
HCN – hydrogen cyanide
HS – HisSer
His – histidine
Im – imidazole
ImpdA – adenosine-5'-imidazole monophosphate
ImpC – cytosine-5'-imidazole monophosphate
ImpdC – deoxycytosine-5'-imidazole monophosphate
ImpG – guanosine-5'-imidazole monophosphate
ImpdG – deoxyguanosine-5'-imidazole monophosphate
ImpT – tyrosine-5'-imidazole monophosphate

ImpU – uridine-5'-imidazole monophosphate
HPLC – high-pressure liquid chromatography
Leu – leucine
Lys – lysine
MALDI – matrix-assisted laser desorption/ionization
MeCN – acetonitrile
MES – 2-(N-morpholino)ethanesulfonic acid
Met – methyl group
MS – mass spectrometry
NMR – nuclear magnetic resonance
NCA – N-carboxyandhydride
ONPG – o-nitrofenylo-β-D-galaktopiranozyd
ONP – o-nitrophenyl-β-D-galactopyranoside
pG – guanine monophosphate
pGpG – guanosine dimer
pGpGpG – guanosine trimer
Ph_3P – triphenylphosphine
Phe – phenylalanine
Pro – proline
RNA – ribonucleic acid
RP HPLC – revered phase HPLC
Ser – serine
SH – SerHis
SIPF – Salt Induced Peptide Synthesis
TFA – trifluoroacetic acid
Thr – threonine
TLC – Thin Layer Chromatography
U – uracil
UV – ultraviolet light
Val – valine

Contents

1. Introduction

1.1. Research on the origin-of-life

Despite the exponential progress in scientific technology which has skyrocketed in the past few decades, research in the field of the origin-of-life has not advanced at the same pace as many other fields in the vast overlapping fields of chemistry and biology. Origin-of-life has not shared this rate of advancement mainly because the problems this particular field faces are not technical but rather of a more conceptual nature. Many questions posed by researchers 30 or 50 years ago are very similar or even exactly the same as the questions we ask today.

The fundamental question of how life could have arisen from an abiotic environment still remains a mystery. The common approach to tackle this problem is to identify the most basic chemical compounds present in all living organisms and determine whether they could be produced under simulated prebiotic conditions.

Prebiotic conditions are defined as; an environment that can reasonably be assumed to be present on the primitive Earth before living organisms arose. Simulated prebiotic reactions utilize simple organic and inorganic molecules such as HCN, CH_4, NH_3, H_2O or H_2 as substrates in the chemical processes that produce more complex molecules, such as amino acids. These simple starting chemicals like HCN or CH_4 are believed to have been present on the early Earth as a result of the formation of solar system. The first experiment of this type was performed in 1953 by Stanley Miller (Miller 1953), in which a mixture of gaseous hydrogen, ammonia, methane and water vapor was passed through electric discharges simulating primordial lightening. Miller obtained a mixture of different amino acids and other complex organic molecules of biological importance. Since then, many similar experiments were performed obtaining various organic and biochemical compounds.

The classic method of prebiotic synthesis developed by Miller works well for thermodynamically stable molecules such as amino acids (Miller 1953, Bada and Lazcano 2003) and nucleobases (Oró 1961, Miyakawa *et al.* 2002), but obtaining biomolecules from inorganic compounds becomes more and more complicated as we climb from nucleosides (Ingar *et al.* 2003) to nucleotides (Powner *et al.* 2009) and is practically impossible for oligonucleotides and polynucleotides (Orgel 2004 and reference therein). The situation is slightly less problematic in case of peptides (Rode 1999), but the general principle of less favorable reactions and smaller yields for

higher molecular weight compounds is present across the board.

The fact that biopolymers, which are essential for living organisms, are not readily produces under assumed prebiotic conditions calls for a great deal of thought when confronting the question of the origin of life. It is very likely that some of the compounds that are essential for modern day life were not freely available in prebiotic times. Instead they were the by-products of preexisting biochemical systems and were included subsequently into the machinery of ascending life. In this scenario, simple substances that are produced during prebiotic synthesis may then have come together to form complex supramolecular structures yielding new reactions and metabolic loops which then produce new compounds previously unachievable (Nelson *et al.* 2000).

It is this very reasoning which lies behind the experimental work here presented: using products which are obtainable under prebiotic conditions and exploiting the properties in order to produce biopolymers that cannot be otherwise produced under assumed prebiotic conditions.

1.2. Prebiotic synthesis of peptides

Amino acids were among the first compounds produced by prebiotic chemistry (Miller 1953). In subsequent experiments many different mixtures of amino acids were obtained under a great variety of conditions (Rode 1999 and reference therein). Not all proteinogenic amino acids are equally easy to obtain in prebiotic reactions. Simple amino acids like glycine, alanine, serine, threonine, aspartic and glutamic acid, proline, valine and leucine are easily obtained under spark discharge experiments (Miller 1993) as well as HCN polymerization (Ferris *et al.* 1978), they are also abundant in examined meteorites (Pizzarello *et al.* 2004), contrary to the more complex amino acids which are rarely obtained in prebiotic reactions; among these are aromatic amino acids, histidine, glutamine and asparagine, methionine, cysteine, arginine and leucine, although somewhat more indirect syntheses have been devised for sulfur-containing amino acids (Van Trump and Miller 1972), histidine (Shen *et al.* 1990a) and aromatic amino acids (Friedman and Miller 1969). It seems, however, very likely that at least some of these “non-prebiotic” amino acids were recruited later, once metabolic/biochemical pathways for other amino acids were established (Cleaves 2009). Glutamine and asparagine are the most obvious candidates for such a scenario (O’Donoghue and Luthey-Schulten 2003).

The next step in “chemical evolution” is the condensation of amino acids to form peptides. Initial attempts in this direction were done already in the 1950's by Sidney Fox (Fox and Harada 1960). In these experiments, a mixture of amino acids (with significant over representation of acidic and basic amino acids) was heated to 160°C for several hours until the complete evaporation of water. Such a procedure was supposed to mimic the “drying lagoon” scenario, in which a mixture of organic compounds from an assumed primordial ocean is hyper-concentrated, thus favoring condensation reactions which require low activity of water to occur (amide bond formation is an example of such a reaction). Fox had observed formation of polymers, but doubted these “protenoids” really were proteins. They were later determined not to be peptides in the classical sense, in fact, the bonds between monomers in those ”protenoids” were more likely a result of interactions between various side chains of amino acids.

The first accepted method of prebiotic synthesis of peptides, performed by Moser *et al.* (1968), was a process in which polymerization and subsequent hydrolysis of HCN yielded a polyglycine chain. This method is however limited only to glycine and cannot be used to obtain heteropolymers.

Also limited to polyglycine were the experiments in which the simulation of a “drying lagoon” was applied to a mixture of amino acids in the presence of various minerals (Lahav *et al.* 1978). This system used the porous surface of certain minerals to pre-organize organic molecules bound to its surface, thus favoring condensation reactions in which cycles of wetting and drying, coupled with temperature fluctuations mimic the natural cycles of tides and daylight, favoring polymerization of glycine.

More recent experimental set-ups for obtaining peptides under prebiotic conditions have succeeded in the synthesis of other homopeptides as well as some heteropeptides, the most accomplished and the best described method from these is the so-called 'Salt Induced Peptide Synthesis (SIPF)' (Schwendinger and Rode 1992). The mixture of amino acids was placed in a solution with high concentration of NaCl and $CuCl_2$ and in temperatures around 85°C, it was observed that short peptides were formed. High concentrations of salt were responsible for coordinating water molecules and coupled with high temperatures to enhance evaporation, the active concentration of water in the system is significantly limited. This allows for condensation reactions otherwise unfavorable in water environments. Additionally, Cu^{2+} ions coordinate amino acids, thus bringing them closer and enhancing the reaction rates (**Fig. 1.1.**). Although

initial experiments were limited to polyglycine (Schwendinger and Rode 1989) subsequent work with this system has shown a wider spectrum of polymerized amino acids (Bujdak and Rode 2004). This method was also successfully combined with the catalytic effects of clay minerals (Rode *et al.* 1999).

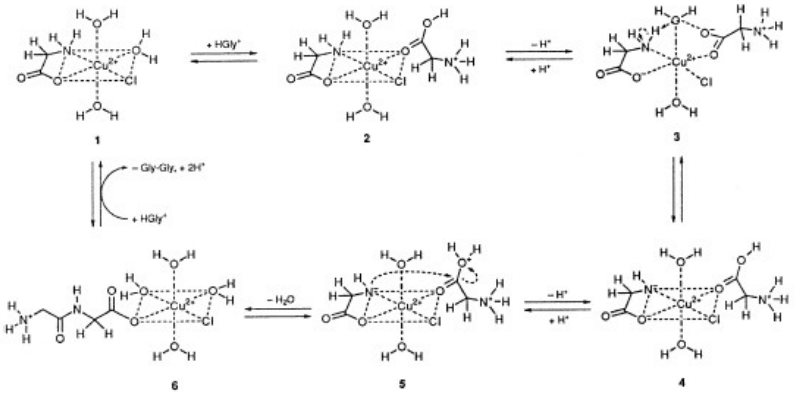


Figure 1.1. Proposed mechanism of SIPF. Gly-Gly example (Rode 1999).

Other methods for obtaining polypeptides coupled drying and rehydratation cycles to microwave radiation as a source of energy (Yanagawa *et al.* 1990). Other, more exotic sources of energy such as ocean floor hydrothermal vents are often proposed as sources for synthesis of organic material on early Earth (Bernstain 2006). Experiments simulating conditions of hydrothermal vents yielded polymers up to octamers of polyglycine starting from a mixture of amino acids (Yanagawa and Kobayashi 1992).

Many other methods of prebiotic synthesis of peptides employ N-Carboxyanhydrate of aminoacids (NCA-aminoacids) at some stage of the synthetic pathway (**Fig. 1.2.**). NCA-aminoacids, also known as Leuch's anhydride, are well known in organic chemistry as precursors for peptide polymerization (Sigma-Aldrich 2010). NCA-aminoacids rapidly hydrolyze in the presence of water, however, if the amount of water is low, NCA-aminoacids preferentially polymerize to peptides (Greenstein and Winitz 1996).

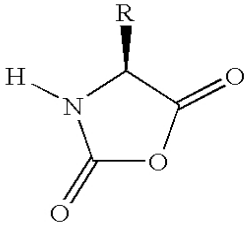


Figure 1.2. The structure of NCA-aminoacid.

Originally, NCA-aminoacids were not considered likely to form in a prebiotic environment. However, various groups have subsequently shown that they can indeed form in plausible prebiotic scenarios, thus giving way to possible formation of peptides.

In some experiments simulating the primordial atmosphere, N-carbamoyl-aminoacids were obtained, while free amino acids were not. (Taillades *et al.* 1998) (**Fig. 1.3.**). Subsequent studies presented a process in which N-carbamoyl-aminoacids in presence of gaseous mixture of NO/O_2 were transformed to NCA-aminoacids which could then polymerize to peptides (Taillades *et al.* 1999)(**Fig. 1.4.**). In addition to polyglycine, various other short chains of heteropeptides were also detected (Commeyras *et al.* 2002).

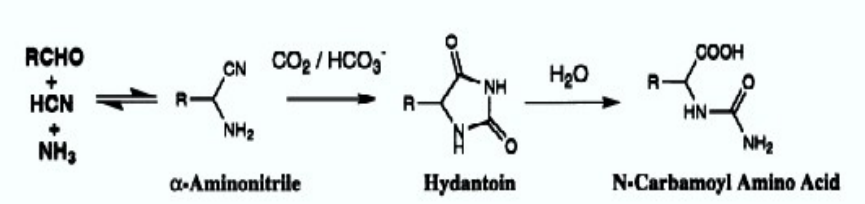


Figure 1.3. The mechanism of the formation of N-carbamoyl-α-aminoacides (Taillades *et al.* 1998).

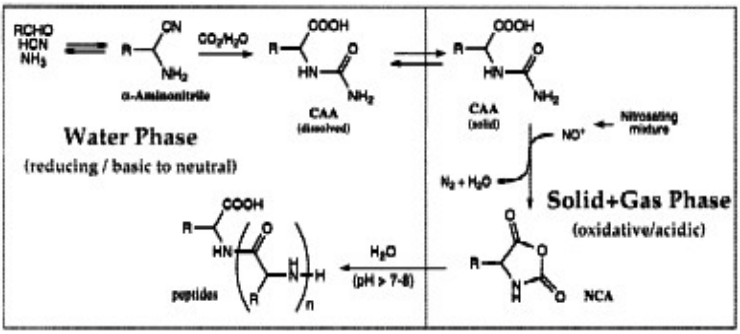


Figure 1.4. The pathway to synthesis of peptides through N-carbamoyl-α-aminoacids and NCA-aminoacids, after Taillades *et al.* 1999.

In a scenario of peptide formation simulating environments influenced by volcanic activity - as a source of carbonyl sulfide (COS) - we also observe an intermediate of N-carbamoyl-α-aminoacids. The presence of this compound can, in favorable conditions, yield NCA-aminoacids, thus paving the way to the formation of peptides (**Fig. 1.5.**). In this scenario it was, however, only possible to synthesize short homo- and heteropeptides (Leman *et al.* 2004).

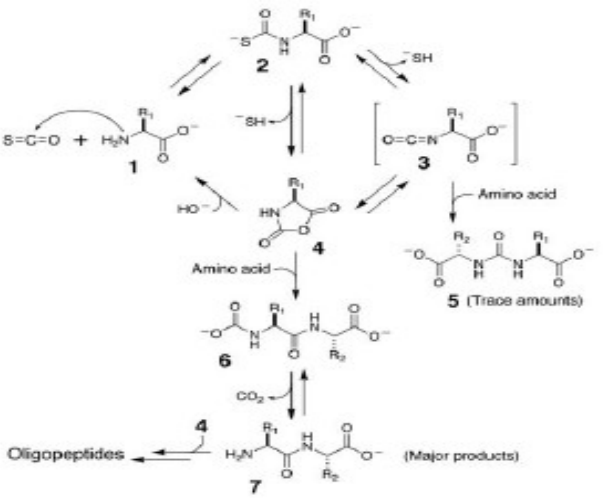


Figure 1.5. The pathway of prebiotic synthesis of peptides through COS and NCA-aminoacids (Leman *et al.* 2004).

Because the concentration of COS in volcanic gasses is not very high it was proposed that other more abundant volcanic gases, namely sulfur dioxide (SO_2), may be used instead of COS as an intermediate for prebiotic synthesis of peptides. Moreover, sulfur dioxide was previously observed in organic chemistry to possess such properties (Kenner and Stedman 1952). Chen and Yang (2007) then proposed a prebiotic scenario in which the synthesis of peptides could occur on the volcanic areas with SO_2 as an intermediate (**Fig. 1.6.**).

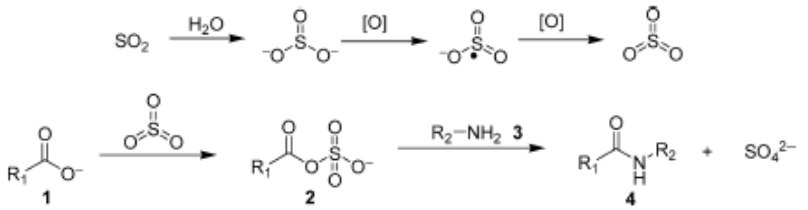


Figure 1.6. The proposed oxidative model for prebiotic synthesis of peptides through SO_2 (adapted from Chen and Yang 2007).

From the above mentioned examples it is clear that conditions that favor the formation of peptides on the prebiotic Earth can be fulfilled, and amino acids can in fact form in many various ways from simpler organic compounds. The presence of amino acids in meteorites (Pizzarello *et al.* 2004, Pizzarello and Shock 2010) provides additional proof that these organic compounds were not unusual in the early Solar System and on premordial earth. The fact that prebiotic simulation experiments produce very similar mixtures of amino acids as those found in meteorites and that the same proteinogenic amino acids are absent in both cases proves that the assumptions behind prebiotic experiments set-ups are well grounded (Cleaves 2009).

The many solutions to condensation of amino acids into peptides argues in favor of the presence of various short peptides in the "prebiotic soup".

1.3. Prebiotic synthesis of nucleic acids

A wide range of biochemical evidence points to the conclusion that the present life, based on information passed from DNA through RNA to proteins, was preceded by some other system in which RNA played a

relatively greater role than in present life forms (Gesteland *et al.* 1999). The evidence for this conclusion stems from a varied array of biological observations; The replication of DNA is initiated by an RNA primer (Itoh and Tomizawa 1980), deoxyribonucleotides are enzymatically synthesized from ribonucleotides (Joyce 1989), the majority of present day coenzymes are ribonucleotides (White 1976), and more importantly, some modern day RNA's catalyze reactions in an analogues manner to enzymes – the so called ribozymes (Kruger *et al.* 1982, Marvin and Engleke 2009).

Countless facts such as those mentioned above have prompted researchers to form the "RNA world" hypothesis (Woese 1967, Gilbert 1986), in which the first stage of life was a single chain of RNA that was able to replicate itself through its activity as a ribozyme, thus unifying genotype and phenotype in one molecule, and solving the chicken and egg problem of molecular biology simply and efficiently.

However, contrary to peptides, the prebiotic formation of nucleic acids has proven itself an arduous, insurmountable puzzle to origin of life researchers.

It seems that nucleic acid bases can form relatively easy; adenine was one of the first products of prebiotic experiments on HCN polymerization (Oro 1961). In a particularly striking experiment, adenine was obtained in 20% yields by heating HCN with liquid ammonia in a sealed tube (Wakamatsu *et al.* 1966); in subsequent experiments, small amounts of guanine were also detected among the products of HCN polymerization (Miyakawa *et al.* 2002).

Some possible reaction pathways from HCN to adenine are presented in **Fig. 1.7.**

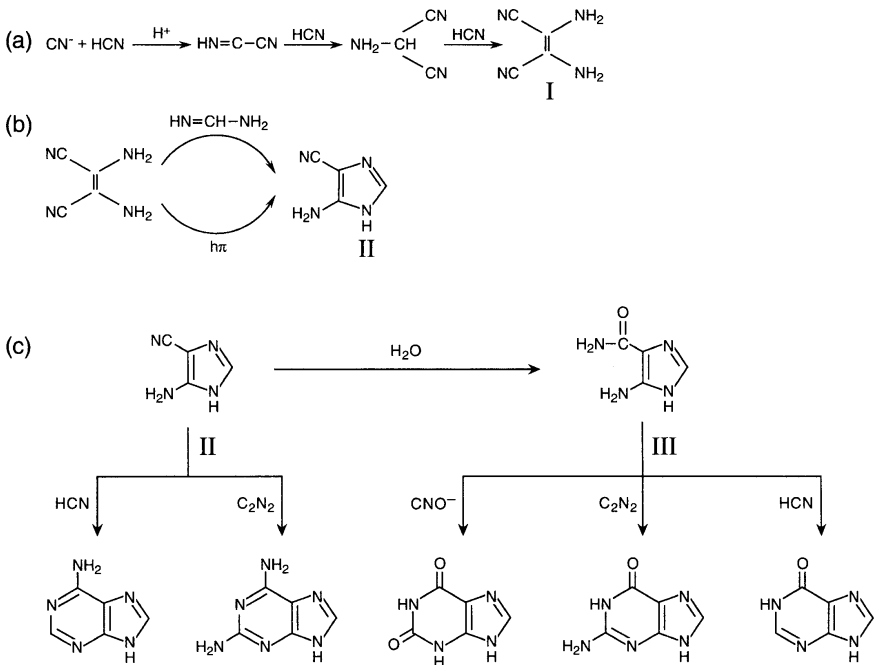


Figure 1.7. Steps in possible prebiotic synthesis of adenine from HCN (Orgel 2004).

Pyrimidine bases are also formed in prebiotic synthesis. Cytosine is obtained in good yields via reactions between cyanoacetylene or its hydrolysis product: cyanoacetalaldehyde and cyanate ions, cyanogen or urea (Ferris *et al.* 1968, Robertson and Miller 1995, Nelson *et al.* 2001): uracil can also be obtained in similar experiments (Robertson and Miller 1995, Bera *et al.* 2010) and was also detected in the Murchinson meteorite (Martins *et al.* 2008).

From the various prebiotic attempts to obtain sugars, the most worthy of note is the Butlerow (formose) reaction. This reaction, which has been studied in considerable detail (Mizuno and Weiss 1974), produces a mixture of sugars in an aqueous environment starting from formaldehyde in the presence of mineral catalysts or heavy metal ions. The Butlerow reaction thus produces a mixture of sugars, from which ribose is usually only a minor product (Decker *et al.* 1982). There are some indications that under certain conditions the reaction can be directed towards aldopentoses, and

more particularly towards ribose (Zubay 1998), however, an efficient, plausibly prebiotic pathway to ribose has not yet been demonstrated.

Even more elusive is the next step in the formation of a nucleic acids: joining a nucleobase to a sugar to form a nucleoside. Attempts to perform such a reaction have had very limited achievements (Fuller *et al.* 1972) and we may safely affirm that, to date, no convincing scenario has of yet been proposed.

The phosphorylation of nucleosides can be achieved prebiotically by incubation with phosphate minerals (Lohrman and Orgel 1968). AMP can be converted to ADP and ATP in the presence of cyanate and insoluble calcium phosphates (Yamagata 1999). However, this fact can be considered of no great value since there is no convincing way to make nucleosides or even ribose.

More promising have been the various attempts to bypass the difficult condensation of ribose and nucleobases and form a nucleoside directly from simpler compounds (**Fig.1.8.**).

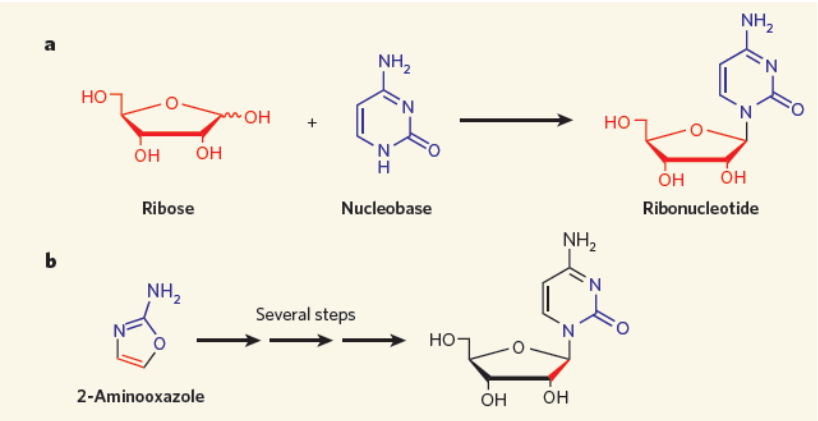


Figure 1.8. Two different approach to obtain a nucleoside (Szostak 2009).

This approach, pioneered in Sutherland's group, still requires a better design to optimize the control of the reactions, but it has, nevertheless, successfully produced β-ribocytidine-2',3'-cyclic phosphate, initially from arabinose-3-phosphate (Ingar *et al.* 2003), and later on from more prebiotic substrates like cyanoacetylene and cyanamide (Powner *et al.* 2009). Researchers in this field have yet to agree whether this approach is prebiotically valid (Szostak 2009), but this work does suggest that the

prebiotic synthesis of at least pyrimidine nucleotides may be possible.

The next logical step in the upwards process of increasing molecular complexity is the formation of polynucleotides, and this too, has proven itself an arduous challenge to researchers; the polymerization of nucleotides in an aqueous solution is a highly unfavorable reaction.

Attempts to polymerize unactivated phosphonucleotides by evaporation and heating met very limited success, forming very short oligomers in which 2'-5'- or 3'-5'-phosphodiester bonds occur more or less at random (Moravek 1967, Rajamani *et al.* 2008). Nucleosides-5'-triphosphates and other nucleoside-5'-polyphosphates, which might seem like an obvious choice of substrates in experiments of non-enzymatic RNA polymerization, did not prove to be effective (Orgel 2004), this is due to the fact that the substrates react so slowly that their polymerization cannot effectively studied in the laboratory.

Thus, nucleotides activated as phosphoimidazolides (**Fig.1.9.**) are commonly used to study non-enzymatic RNA polymerization.

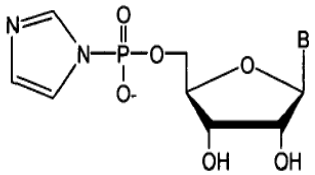


Figure 1.9. Phosphoimidazole derivative of ribonucleotide.

Phosphoimidazolides are soluble in water, easily hydrolyzed and, in favorable conditions, can polymerize, yielding RNA chains with various degrees of natural 3'-5' and unnatural 2'-5' linkages. Despite some claims (Lohrmann 1977) these substrates are not generally considered to be prebiotically available, they are wildly used as they might give us an insight on how polymerization could occur in a non-enzymatic fashion.

Highly concentrated phosphoimidazolides can polymerize to short oligomers in an aqueous solution and the introduction of various catalysts to the system can yield more effective polymerization. Particularly metal ions like Pb^{2+}, Mg^{2+} an Mn^{2+} can be used to catalyze this reaction (Sawai and Orgel 1975, Joyce 1987).

Using an eutectic solution of water-ice, which up-concentrate the nucleotides, together with metal ion catalysis can yield even longer

oligomers and of every base (uracil derivatives are the least active and in some methods are not effectively polymerized) (Kanavariotti *et al.* 2001, Monnard and Szostak 2008).

A great deal of attention has been given to the catalysis of polymerization by montmorillonite - a clay mineral. It is a mineral with layered aluminosilicate structure that expands to permit large molecules to enter the inter-layers. Oligomers up to 40mer were synthesized using montmorillonite as a catalytic substrate (Huang and Ferris 2003). In certain settings, a regioselectivity of 80% towards 3'-5' bonds could be observed (Prabahar and Ferris 1997). Such experiments are of great interest as mineral catalysis is likely under prebiotic conditions.

Other studies using phosphoimidazolides concentrated on exploring the kinetics of template directed, non-enzymatic polymerization. Various systems have been tested, including hairpin primer elongation (Hill *et al.* 1993, Monnard and Szostak 2008).

From studies of template directed synthesis and from other approaches we know that it is indeed possible to obtain a nucleic acid chain in an non-enzymatic environment, starting from single nucleotides, providing that they are properly activated. However, worthy of note is the fact that in none of the systems a full replication cycle can been obtained – a template in order to be replicated non-enzymatically needs to have at least 60% of C residues, therefore disabling formation of complementary template (Orgel 2004). The kinetics of most non-enzymatic methods have similar relationships; incorporation of a G residue opposite a C is quite efficient, while incorporation of an A opposite a U and a C opposite G is less efficient, and the incorporation of a U opposite an A is next to impossible (Orgel 2004). A pair of consecutive A residues in a template almost completely blocks further synthesis.

This, coupled with the fact that no prebiotic synthesis can produce activated nucleotides that could be used as substrates for further oligomerization provides evidence not only in opposition to the RNA world hypothesis but also provides the general terms to argue the presence of RNA or any significant amount of nucleotides at the earliest stages of chemical evolution.

1.4. Peptides as prebiotic catalysts

Modern life uses enzymes to execute its functions; enzymes are long chain polypeptides that, thanks to unique folding, can bring together side residues of amino acids into precise orientations so that they may catalyze given reactions. We know that peptides can be formed under prebiotic conditions, however, all these processes tend to make short peptides and even if they were able to make longer ones it is not clear how, in a prebiotic scenario, we could make many copies of the same long sequence (Luisi 2006, 2007).

A highly relevant question which follows in light of the above mentioned observations is the following: Can short peptides have enzyme-like activities? Could peptide catalysis be a determining factor in the prebiotic world? Although for many years it seemed like macromolecular folding is a necessary prerequisite for peptide catalysis (Luisi 1979), many examples of short peptides being catalytic factors in various reactions have recently been given in experimental literature .

The early reports on such possibilities were purely phenomenological in nature. In some simulations of prebiotic reactions it was observed that the dipeptide His-His, which formed in those reactions, enhanced the yields of subsequent peptide condensations (White and Erickson 1980, Shen *et al.* 1990b).

Similar cases were described for Gly-Gly, which is formed as a product in Salt Induced Peptide Synthesis (SIPF); once it is formed, it enhances the rate of subsequent peptide condensations (Plankensteiner *et al.* 2002).

The oligopeptide $[Leu\text{-}Lys]_n$ ($n \geq 10$) was reported to catalyze the hydrolysis of nucleic acids (Barbier and Brack 1987); it was shown that these properties depend on the regularity of exposed basic side chains (Barbier and Brack 1992).

A metallo-dipeptide, Cys2-Fe^{2+}, was reported to catalyze hydrolysis of o-nitrophenyl-β-D-galactopyranoside (ONPG), a standard substrate for detecting activity of β-galactosidases. In anhydrous conditions it was possible to reverse the reaction and demonstrate condensation of ONPG from its usual hydrolytic products: β-D-galactose and o-nitrophenol (ONP). Since it was also reported that OMPG can assist in formation of Cys2-Fe^{2+} from Cys-Cys and Fe^{2+}, the possibility of an autocatalytic cycle involving a short peptide came to light, something of great prebiotic significance

(Fleminger *et al.* 2005).

Another example of the catalytic activities of small peptides is an aldol reaction. It has been known for quite some time that proline can be used as an enantiomeric catalyst of the aldol reaction (List *et al.* 2000). It was proven that not only proline, but also peptides with proline at their N-end can show similar properties (for example Pro-Gly, Pro-Glu-Leu-Phe) (Kofoed *et al.* 2003). Subsequently it was also shown that short peptides with primary amine at the N-end can catalyze aldol reactions with good yields and enantioselectivity (for example Ala-Ala, Val-Val, Val-Phe, Ala-Ala-Ala) (Zou *et al.* 2005). Although these were reactions with relevance to organic chemistry, it was later possible to demonstrate the same process in water environments, thus compatible with prebiotic scenarios (Dziedzic *et al.* 2006).

This later experiment led other researchers to use the Val-Val dipeptide as a catalyst for enantiomeric enhancement in tetrose synthesis (Weber and Pizzarello 2006). Since it is known that meteorites have relatively greater percentage of amino acid L enantiomers then their D enantiomers (Cronin and Pizzarello 1997), it was proposed that those aminoacides could be the base for the passage of enantiomeric enhancement on early Earth *via* the aldol reaction (Pizzarello and Weber 2004). It was also shown that the aldol reaction catalyzed by L-Val-L-Val had an 80% enantiomeric excess of D enantiomer of synthesized erythrose (Weber and Pizzarello 2006).

From an origin-of-life perspective, an important discovery was that of the catalytic properties of the dipeptide SerHis (**Fig.1.10.**). This dipeptide was reported to possess broad hydrolytic activities: SerHis could hydrolyse esters as well as proteins and nucleic acids (Li *et al.* 2000), in nature many of the reactions catalyzed by SerHis have specialized enzymes, of which most are serine hydrolases. SerHis is also a common active group in many of the most important groups of enzymes, including chymotrypsin, subtilisin and Phospholipase A2, comprising approximately 1% of the genes in the human proteome (Simon and Cravatt 2010). The active sites of these enzymes are composed of the side chains of the amino acids serine, histidine and aspartic acid. This led to speculation that the dipeptide SerHis could be a primitive analog of the serine proteases.

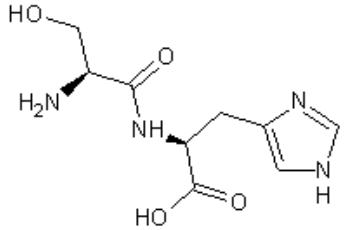


Figure 1.10. The structure of SerHis dipeptide.

The mechanism of serine hydrolases is well understood (**Fig.1.11.**). In the case of peptide hydrolysis (as an example), the -OH group of serine acts as a nucleophile, attacking the carbonyl carbon of the peptide bond to be cleaved. The free electron pair from the imidazole histidine ring can accept hydrogen from an -OH group, thus coordinating the attack of the peptide bond. The electronegativity of the free electron pair in the imidazole is greatly enhanced by the carboxyl group from the aspartic acid, which, by hydrogen bonding, binds the imidazole ring (Kraut 1977).

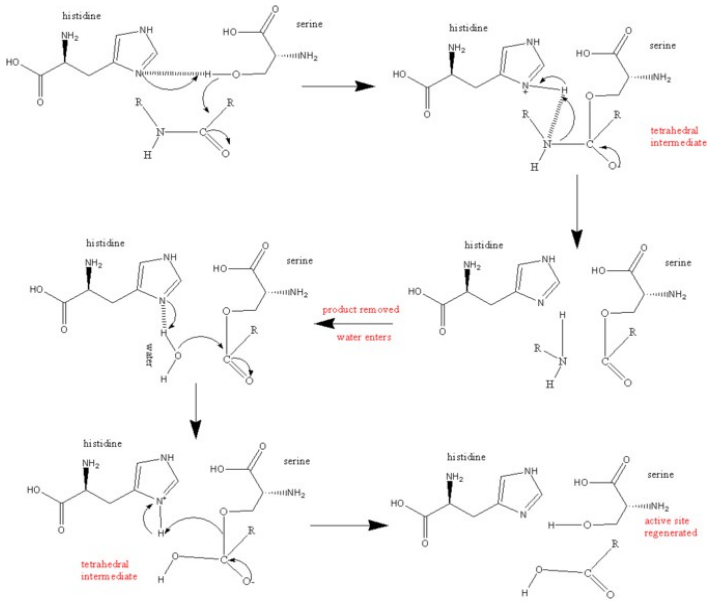


Figure 1.11. The active site of serine proteases. Involvement of serine and histidine side chains (Wikipedia).

These insights into the mechanism of serine proteases (and other serine hydrolases) can be very useful in understanding the mechanism of the dipeptide SerHis catalysis.

Other peptides with similar structures to SerHis were also investigated (Li *et al.* 2000). Some of these were inactive, others had smaller or equal levels of activity, however, none other dipeptide exceeded the catalytic capability of SerHis (these dipeptides are presented in **Table 1.**).

Table 1. The proteolytic activity of various short peptides. Catalytic activity relative to SerHis is indicated (adapted from Li *et al.* 2000).

amino acid or peptide	hydrolysis of proteins
Ser	0 %
His	0 %
Ser + His	0 %
HisSer	0 %
SerHis	100 %
SerHisAsp	100 %
CysHis	60 %
ThrHis	20 %
GlySerHis	20 %
SerHisGly	60 %
SerHisHis	60 %
SerGlyHisHis	40 %
SerGlyGlyHisHis	40 %

From a prebiotic perspective, the desired catalytic activity would be not to hydrolysis of these compounds but rather their synthesis; but hydrolysis and condensation are two sides of the same coin (**Fig.1.12.**). The direction of the reaction, namely hydrolisis or condensation, depends on the thermodynamics, the favorable direction will change in response to the ratio between the reactants. A catalyst for this reaction will catalyze either

hydrolysis or condensation depending on the conditions. This is why it is possible to use proteolytic enzymes to synthesize peptide bonds instead of peptide bond cleavage (Schuster *et al.* 1990, Bordusa 2002).

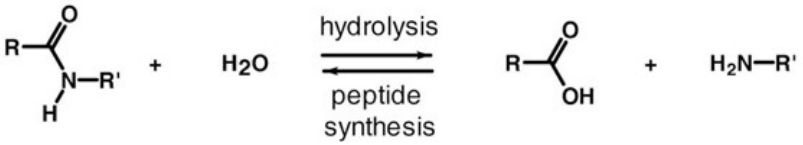


Figure 1.12. Hydrolysis and synthesis of peptides are two directions of the same reaction (from Gorlero *et al.* 2009).

It is well known that SerHis catalyzes the hydrolysis of peptide bonds (Li *et al.* 2000, Chen *et al.* 2001, Du *et al.* 2002). Can therefore this activity be reversed, so that SerHis will catalyse the formation of peptide bond? It was shown that this is indeed possible (Gorlero *et al.* 2009). In a situation where two amino acid substrates are soluble in the reaction mixture and the peptide product is insoluble, the peptide will drop out of the solution, thus driving the chemical equilibrium towards synthesis. SerHis catalyzing this reaction therefore catalyzes the formation of peptides (Gorlero *et al.* 2009).

It was shown that even short peptides can have diverse catalytic properties. This is very important from the origin-of-life perspective. Since short peptides can be formed under prebiotic conditions and short peptides do in fact have catalytic activity, it can be assumed that peptide catalysis was a relevant factor at the time of the "prebiotic soup". This is particularly important for the catalysis of the formation of peptide bonds shown by the dipeptide SerHis.

Geochemistry can produce compounds of only limited complexity. Where synthesis based on purely abiotic geochemical processes has stopped, a complex, bio-organic process could pick up the synthesis, taking molecular complexity to a higher level.

1.5. Aim of the PhD project

Amino acids and peptides form easily under prebiotic conditions. In contrast, nucleotides and nucleic acids are not obtainable under such conditions. Certain short peptides can catalyze various reactions, most notably, hydrolysis of diverse compounds; hydrolysis is a reaction that can be reversed if conditions are favorable. The SerHis dipeptide can efficiently catalyze the hydrolysis of other peptides (Chen *et al.* 2001, Du *et al.* 2002) and this reaction can be reversed to catalyze the condensation of peptides (Gorlero *et al.* 2009). SerHis can also catalyze the hydrolysis of nucleic acids (Li *et al.* 2000, Ma *et al.* 2007); is it then possible to reverse this reaction and achieve a prebioticaly plausible peptide catalyzed synthesis of nucleic acids, for example of RNA (**Fig.1.13.**)?

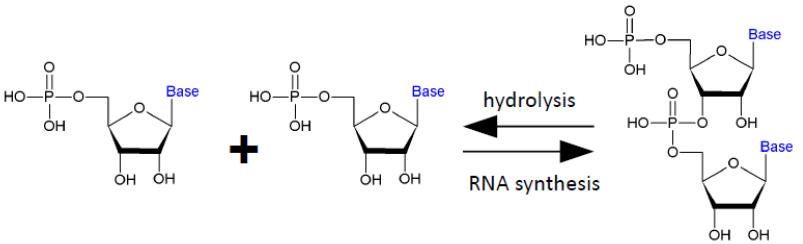


Figure 1.13. Hydrolysis and synthesis of nucleic acids are two directions of the same reaction.

Such an achievement would be of significant importance for origin-of-life research, it would mean that some stumbling blocks of prebiotic synthesis can be removed by catalytic activities embodied in simpler products of prebiotic synthesis.

The aim of this work is to explore such a possibility, investigate the conditions required to reverse this reaction, and explore the properties of the catalyst as well as the mechanism of catalysis.

2. Results

2.1. Hydrolysis of nucleic acids by SerHis

To successfully reverse the hydrolytic activity of SerHis we first need to better understand its phosphodiesterase properties. An experiment demonstrating the pH dependance of phosphodiesterase activity of SerHis is presented in **Figure 2.1.** Prolonged incubation of a linear, double stranded DNA fragment in the presence of SerHis resulted in a smear of digested DNA. The digestion was present only at pH 6 and below. In higher pHs, the dipeptide does not exhibit phosphodiesterase activity. The imidazole sidechain of histidine has a pK_a of approximately 6.0, this indicates that SerHis, in order to hydrolyze nucleic acids, has to have its imidazole ring protonated.

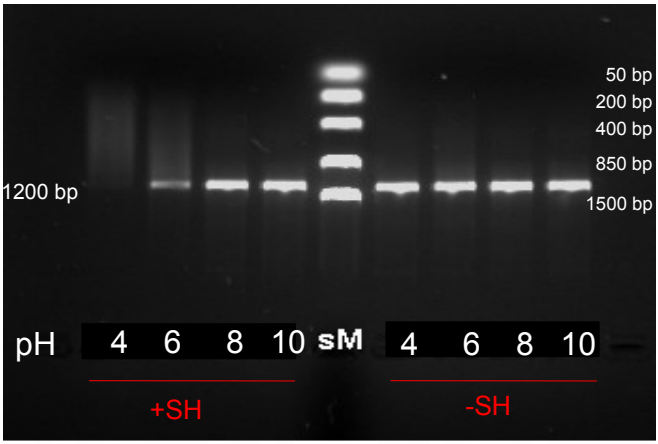


Figure 2.1. Hydrolysis of 1200 bp 30 ng/µl linear dsDNA by 8 mM SerHis, T=37°C, time=5days, 40 mM Britton-Robinson buffer pH 4, 6, 8, 10.

The properties of the SerHis dipeptide are highly specific. Neither the dipeptide HisSer, nor the single or combined amino acids Ser and His possess clear hydrolytic activities (**Fig.2.2.**). Under extreme conditions (48 hours in pH=4,0 and 45°C), only histidine containing amples, can show

some enhancement in hydrolysis over a control sample. However, these activities were negligible in comparison with SerHis, which in the above mentioned conditions completely digested the DNA sample.

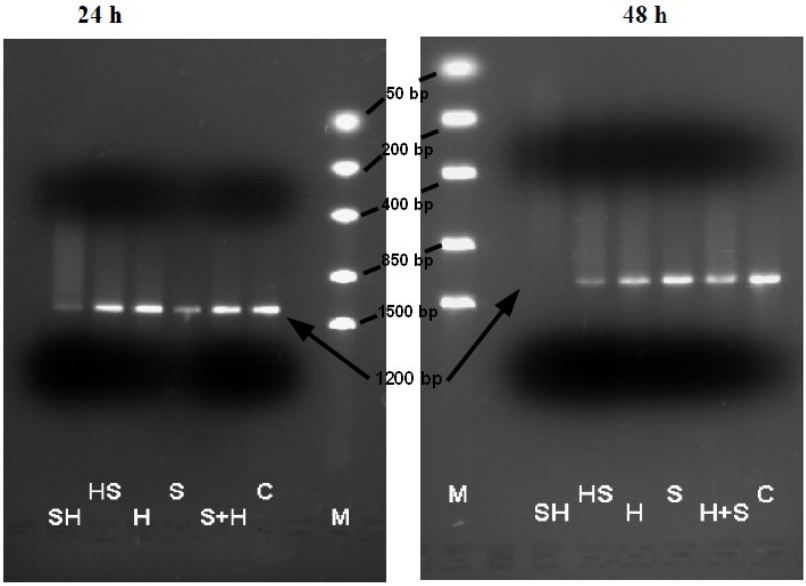


Figure 2.2. Hydrolysis of 1200 bp linear dsDNA by different SerHis related compounds. 30 ng/µl l DNA, 8 mM of tested compounds. T=45°C, time=24 and 48 hours, 40 mM Britton-Robinson buffer pH 4,0. SH – SerHis dipeptide, HS – HisSer dipeptide, H – histidine, S – serine, H+S – histidine plus serine, C – control, no compound. In a gel after 24 hours, half of the S (serine) sample was spilled resulting in smaller band. Serine exhibited no hydrolytic activity after 48 h incubation. Thus, the apparent degradation after 24 h was an artefact.

This result clearly shows that the properties of SerHis originate from the specific arrangement of its chemically active groups (amino group of N-end, carboxyl group of C-end, hydroxyl group of serine and imidazole ring of histidine) and not because of the sole presence of those groups.

This is a property that SerHis shares with enzymes; catalytic activity depends on the structural arrangement of the sidechains.

2.2. Synthesis of nucleic acids by SerHis

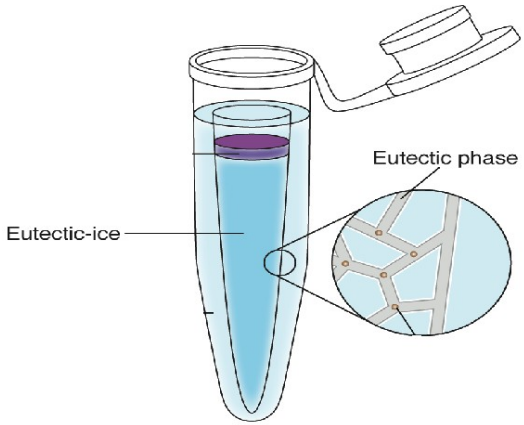
2.2.1. Introduction

The reaction environment used in this work is the eutectic phase in water-ice. There are several reasons for which this environment was chosen. The main difference between classical synthetic organic chemistry and prebiotic chemistry is that the later requires the use of a water environment. However, in water, due to entropic reasons, the hydrolysis of polymers, such as nucleic acids, is favored over condensation. For this reason, various tricks are employed in prebiotic chemistry to lower the chemical activity of water in nominally aqueous environments (see chapter 1). Freezing water below 0°C is one of those "tricks".

Water/ice systems are formed when aqueous solutions are cooled below their freezing point but above the eutectic point. In these environments, the two phases coexist and form the eutectic phase system: a solid (made of pure water ice) and a liquid phase containing most solutes originally present in the whole sample. The exact temperature of the eutectic point depends on the concentrations of various molecules, like buffer compounds or inorganic ions, in the sample, but in most of the employed systems it falls slightly below -20°C (Monnard and Ziock 2008).

During freezing, the ice begins to nucleate from solution, forming pure ice. If the eutectic point is not achieved the ice crystals will grow to certain size but the whole sample will not freeze completely, thus leaving liquid microchannels from the remaining water and other compounds present in the sample (those compounds are highly upconcentrated) (**Fig. 2.3.**).

This eutectic environment presents three major advantages for polymerization reactions: i) such an environment can very efficiently concentrate solutes and likely allows for their self-assembly into organized structures such as stacks, which are required for efficient polymerization ; ii) freezing reduces the water activity in the system by dehydrating the sample, and iii) due to the low temperatures, activated monomers and RNA products are protected against thermal decomposition (Monnard and Ziock 2008). In addition, the medium has been shown to favor the non-enzymatic, metal-ion catalyzed polymerization of various activated ribonucleotides.



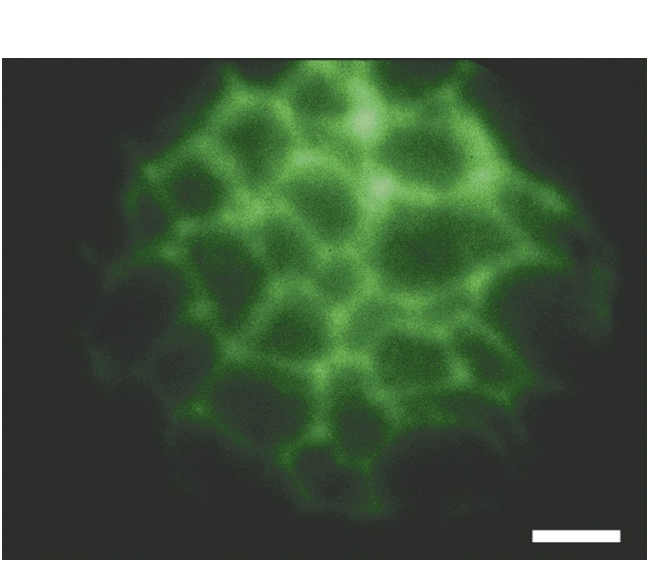


Figure 2.3. The structure of water/ice eutectic system. **A** : eutectic phase forming in an Eppendorf tube placed in below zero temperature (adapted from Attwater *et al.* 2010). **B** : Epifluorescence micrograph of partly frozen nucleotide monomer suspension used in self condensation experiments. Acridine Orange was added to visualize the structures, sizebar = 27.4 μm (Kanavarioti et al 2001).

In the standard conditions used throughout this thesis (described below) SerHis had only slightly elevated the decomposition rates of RNA (**Fig.2.4.**).

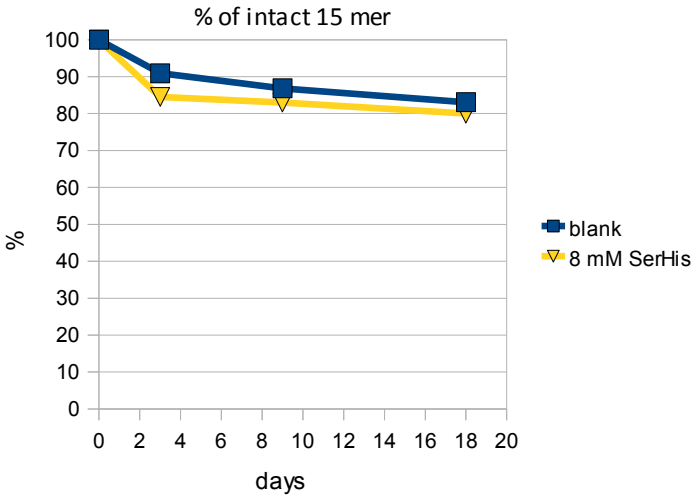


Figure 2.4. Degradation of single strand 15 mer RNA under standard conditions.

Low temperatures and low activity of water in the eutectic water/ice system resulted in unfavorable conditions for hydrolysis. Thus over the period of 20 days and in the presence of otherwise hydrolytic peptide SerHis more than 80% of RNA was intact. Conditions of suppressed hydrolysis should in principle be compatible with condensation reactions.

2.2.2. Results

Similarly to other non-enzymatic systems, the polymerization of monomers of nucleic acids with SerHis as a catalyst was achievable only with the introduction of activating groups. The imidazole activating system is the most common in those type of experiments (Orgel 2004). ImpG – guanosine 5'-phosphoimidazolide (**Fig.2.5.**) was used as a principal substrate in polymerization experiments.

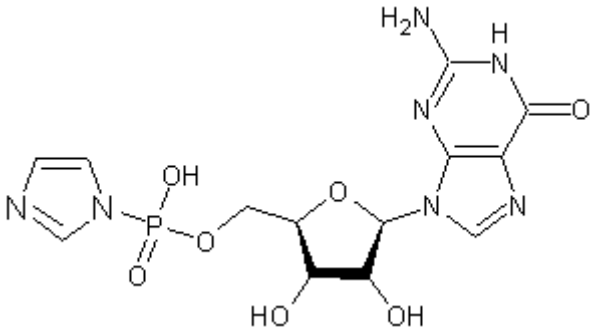


Figure 2.5. The structure of guanosine 5'-phosphoimidazolide – ImpG.

ImpG was synthesized as a sodium salt using the method described by Kanavarioti *et al.* In 1999, with the purity of 98±1%. The quality of the synthesis was determined by the use of HPLC and MALDI-MS (see Materials and Methods for description).

The standard conditions of the majority of the reactions are presented in **Figure 2.6.** A 100 µl water solution containing 5 mM ImpG, 4 or 8 mM SerHis, 5 mM MES buffer pH 6,5 was prepared and placed in a ethylene glycol:water (1:1) bath kept at -18,4°C. Samples were incubated for up to a month time. The processes of the reactions were monitored by a diode-array UV-VIS spectrophotometer coupled with ion-exchange HPLC column (see Materials and Methods chapter for more details).

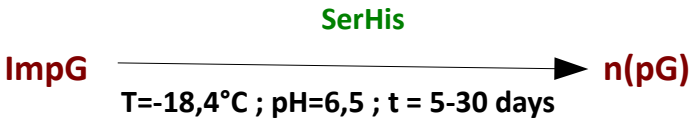


Figure 2.6. The standard conditions of SerHis catalyzed polymerization of ImpG.

Under standard conditions ImpG polymerized to short oligomers. In the eutectic phase polymerization similarly to other non-enzymatic polymerizations we expect great variety of products. Except of natural 3'-5' phosphodiester bonds we also encounter 2'-5' bonds as well as

pyrophosphate dimer and pyrophosphate capped oligomers with 2'-5' and/or 3'-5' bonds (Monnard *et al.* 2003).

Within 10 days time around 10% of initial ImpG was polymerized to oligo(G). This figure includes all types of dimers and longer oligomers with the exception of pyrophosphate dimer (it also forms in a not catalytic pathway). The figure remained stable during 1 month of monitoring (**Fig.2.7.**).

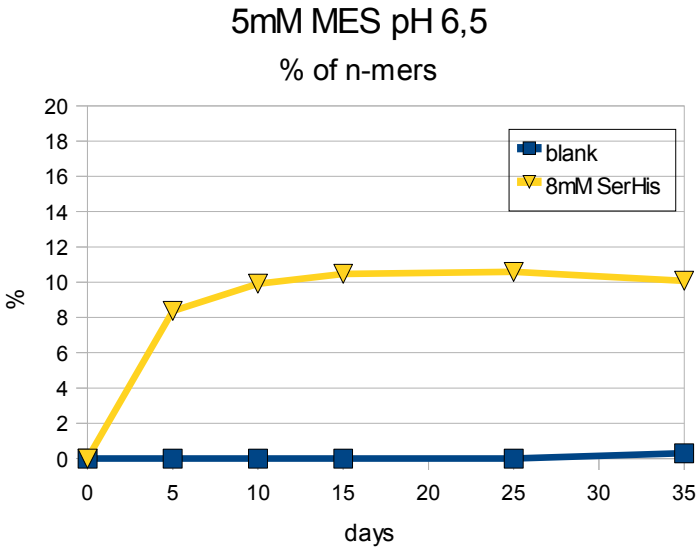


Figure 2.7. Yield of oligo(G) synthesis under standard conditions.

ImpG polymerization was observed under all tested pH conditions in the range from 5,0 to 8,2 (**Fig.2.8.** and **Fig.2.9.**). Maximum yield of oligomers was 10%. This yield decreased in pH below 6, presumably due to enhanced hydrolysis of the substrate. In pHs above 6, however, the yield remained constant, indicating that the rate of deprotonation of the histidine's imidazole ring is not the decisive factor in polymerization rates.

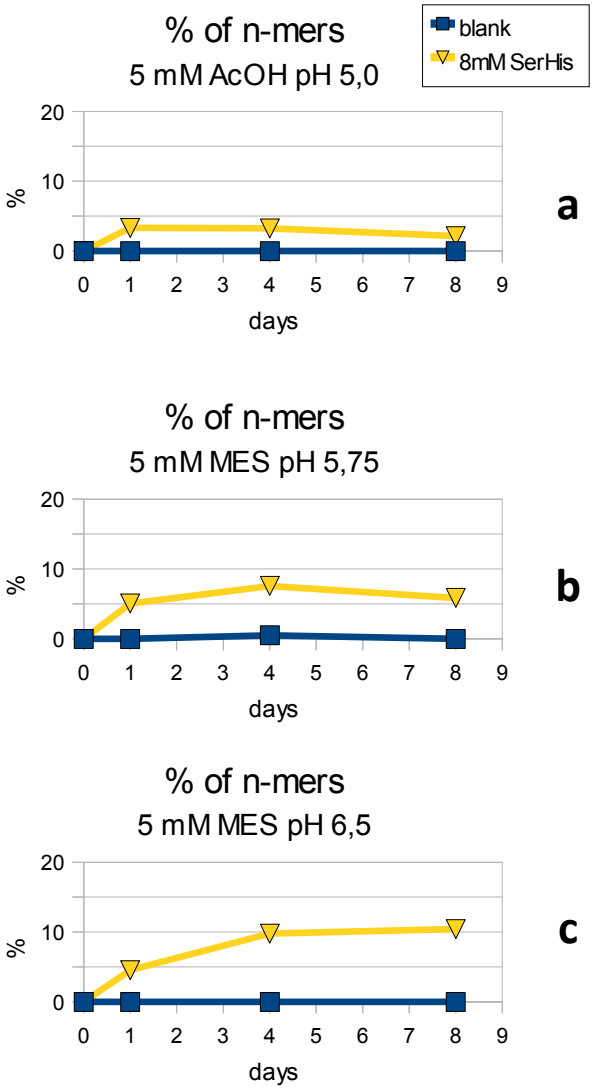


Figure 2.8. Yield of oligo(G) synthesis under low pH conditions. a – 5mM AcOH buffer pH5,0 ; b – 5mM MES buffer pH5,75 ; c – 5mM MES buffer pH6,5.

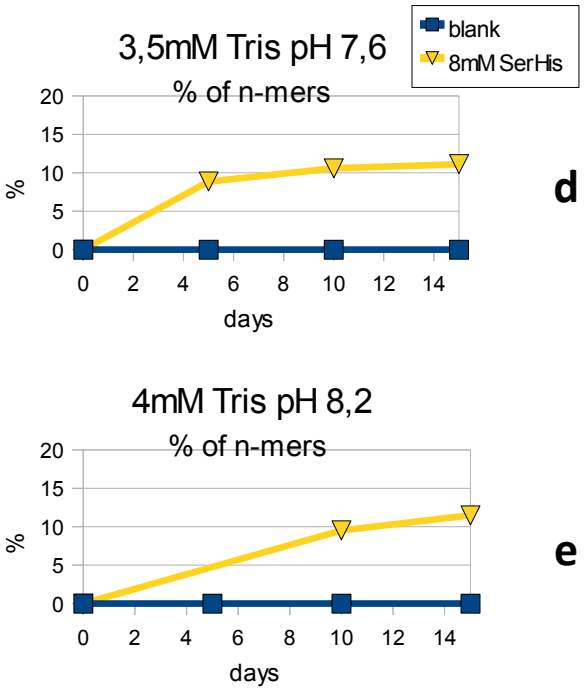


Figure 2.9. Yield of oligo(G) synthesis under basic pH conditions. d – 3,5mM Tris buffer pH7,6 ; e – 4mM Tris buffer pH8,2.

Polymerization was observed in all three different buffers tested, showing that the activity of SerHis is not greatly influenced by buffers, which was one of the concerns surrounding catalytic activity of SerHis (Sun *et al.* 2004).

As with the ability to hydrolyze nucleic acids (**Fig.2.2.**), the ability to synthesize nucleic acids was also limited highly specific to the SerHis dipeptide (**Fig.2.10.**). Other related compounds, including the isomeric dipeptide HisSer as well as mixture of serine and histidine, did not show any activity. Once again, this demonstrated that the catalytic activity of SerHis stems from enzyme-like geometrical arrangement.

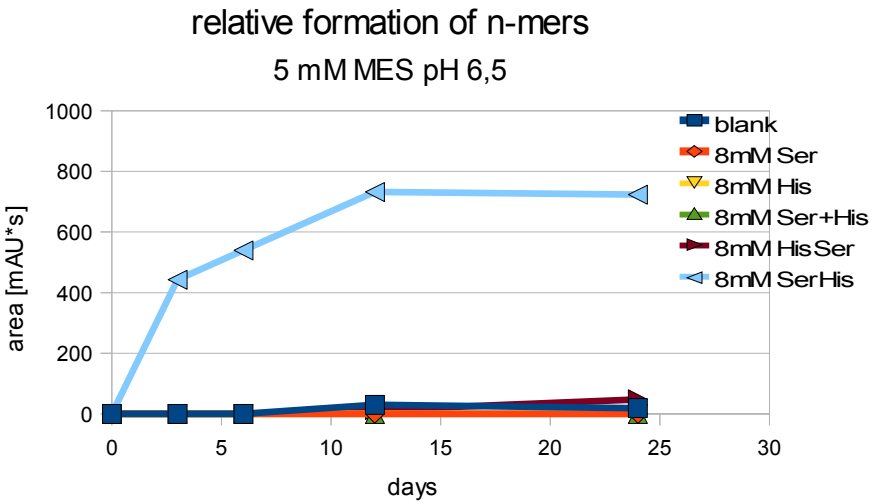


Figure 2.10. Oligo(G) synthesis by various SerHis related compounds in standard conditions.

The imidazole leaving group is the most common of the leaving groups employed in non-enzymatic polymerization of nucleic acids. However, various other derivatives of imidazole are also used as activating groups. To test whether the activity of SerHis is limited by the type of activating group, different derivatives of ImpG were synthesized and tested for SerHis catalyzed polymerization (**Fig.2.11.**). The relative rate of polymerization of mononucleotides with different activating groups is well studied (Monnard *et al.* 2001, Huang and Ferris 2003). Different yields presented in **Fig.2.11.** are in agreement with previously published data. An important factor for the peptide catalysis thesis is that SerHis interacts with every substrates, therefore its activity is not limited to one activating group.

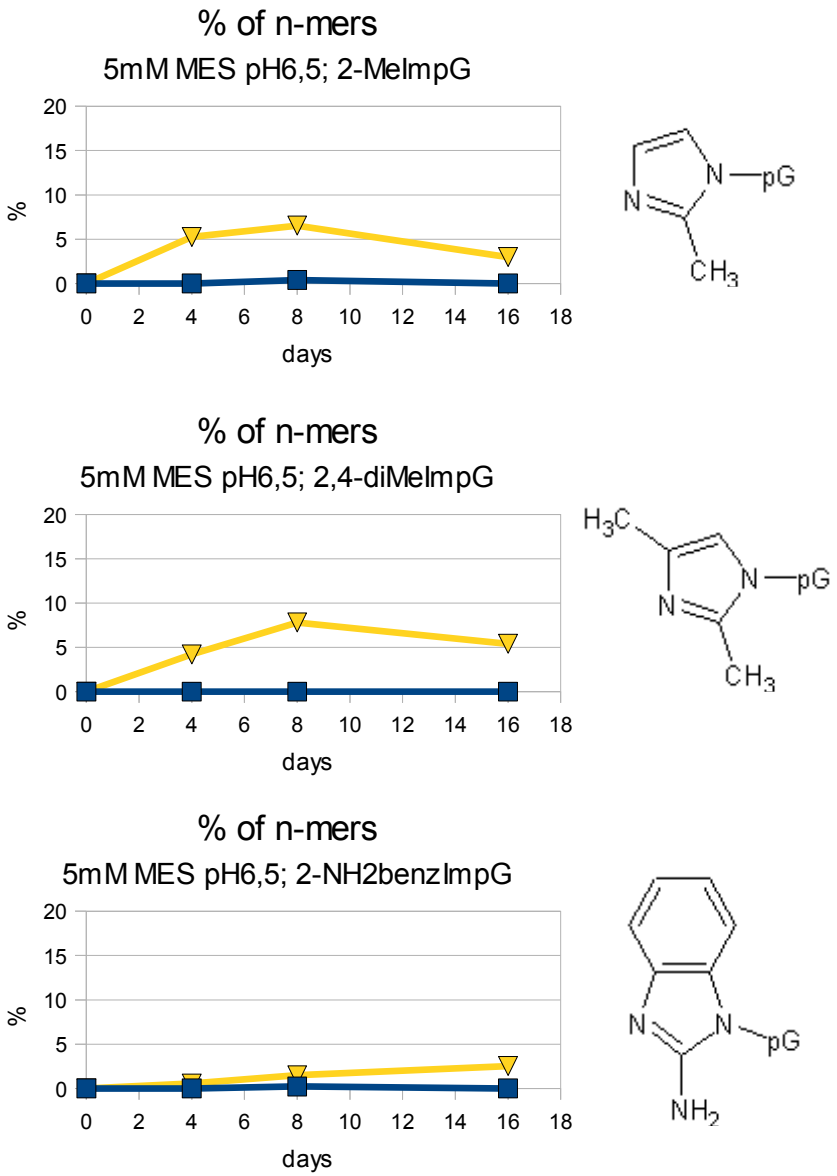


Figure 2.11. Yield of oligo(G) synthesis with different activating groups. Top – 2-MeImpG ; midlle – 2,4-diMeImpG ; bottom - 2-NH_2-BenzImpG.

The next factor that needed to be determined was whether or not the observed activity of SerHis was limited only to guanine nucleotides or whether it would react in a similar manner with other canonical nucleotides.

As can be seen in **Fig.2.12.**, dimers of all three remaining nucleobases were obtained from activated mononucleotides. SerHis was a catalyst in all three cases, proving its universal properties.

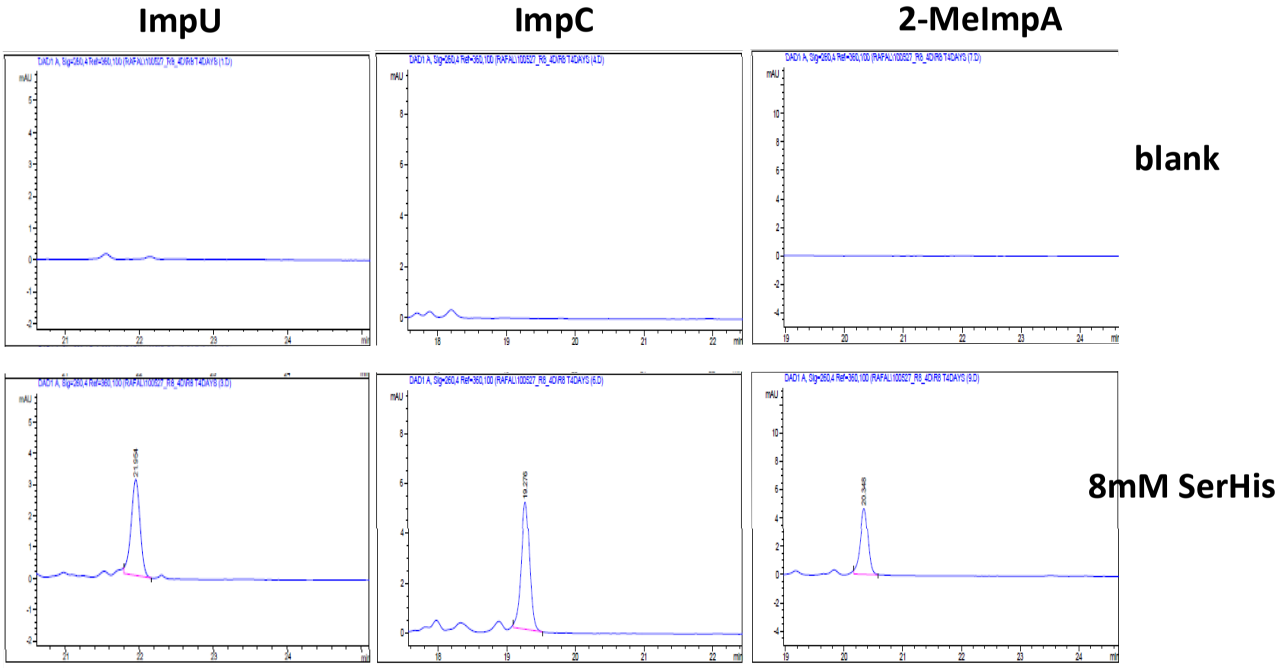


Figure 2.12. Chromatograms of dimeric regions of polymerization of activated monomers of different RNA bases. Ser-His catalyzed polymerizaion in each case.

Besides the RNA nucleotides, imidazole activated DNA nucleotides were also tested (**Fig.2.13.**). In every case, with the exception of deoxycytidine, the presence of SerHis had catalytic effects on the formation of dimers. It is established that in non-enzymatic nucleic acid polymerization, the 2'-OH groups are more reactive than 3'-OH groups (Kanavarioti *et al.* 1999). This is another reason why most of these types of experiments deal with ribonucleotides and not with deoxyribonucleotides and why, lately, 3'-deoxy artificial nucleotides are being explored (Schrum *et al.* 2009). The ability of SerHis to work with rybose as well as deoxyrybose is of prime interest.

In addition to various imidazole activated mononucleotides, also non activated nucleotides and 5'-triphospho activated nucleotides, were tested. None of them showed any significant polymerization. In the standard employed conditions, ATP was much more stable than ImpA. Imidazole activated mononucleotides were slowly being hydrolyzed, whereas ATP remained almost unchanged during one month of incubation.

The highest yields of nucleotide oligomerization was achieved with SerHis concentrations close to equal to that of the ImpG substrate (**Fig.2.14.** and **Fig.2.15.**). SerHis is sold as a non-stechiometric salt of acetic acid and water (Bachem 2009) meaning that the actual concentration of SerHis is only around 70% of its the nominal, thus 8 mM SerHis is approximately equal to 5 mM ImpG.

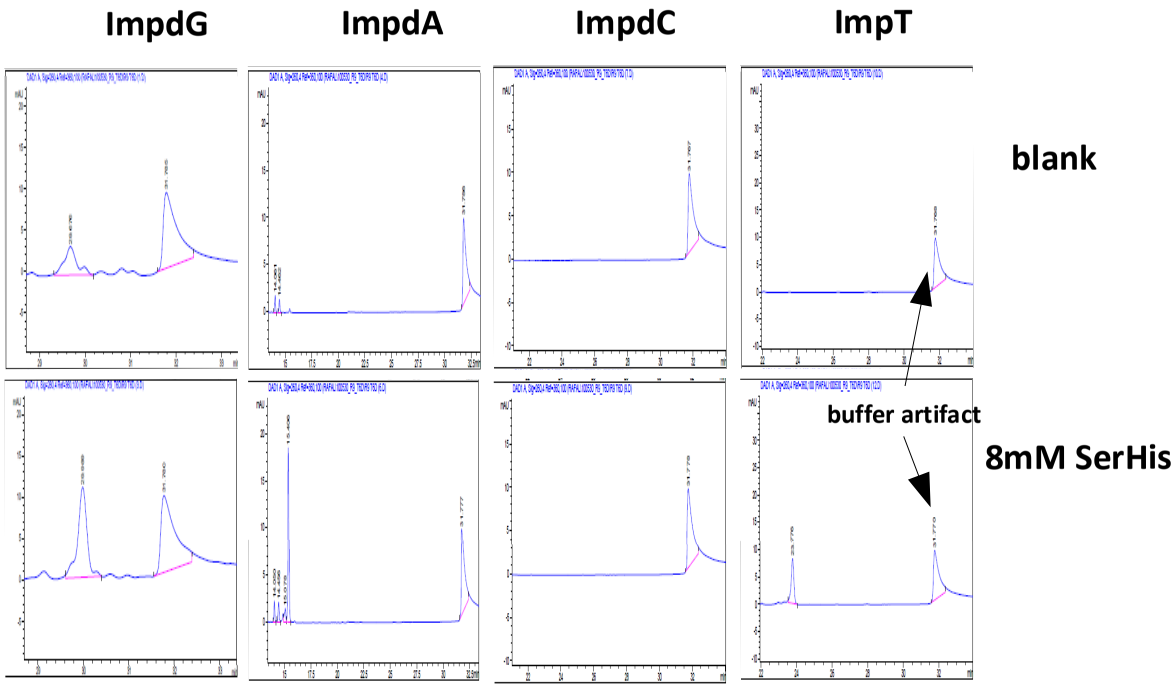


Figure 2.13. Chromatograms of dimeric regions of polymerization reactions of different bases of activated DNA monomers.

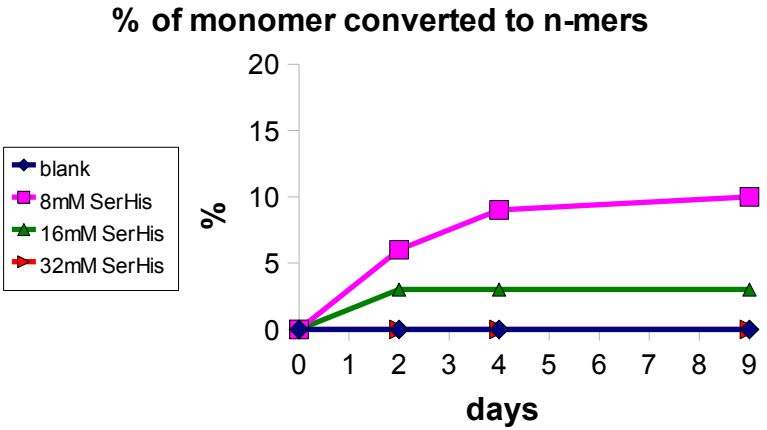


Figure 2.14. Yield of oligo(G) synthesis with different concentrations of SerHis under standard conditions.

Higher concentrations of SerHis gave lower yields of oligomerization presumably because of greater hydrolysis of ImpG. In lower concentrations yields drop proportionally to the amount of SerHis (**Fig.2.15.**).

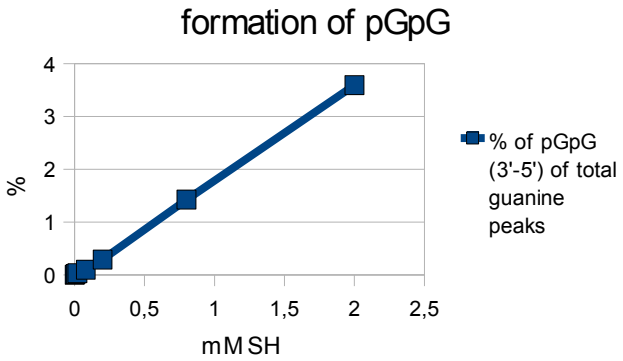


Figure 2.15. Yield of 3'-5' guanine dimer formation with different concentrations of SerHis under standard conditions after 10 days and 25mM ImpG as starting substrate.

With the reaction of condensation of ImpG into a 3'-5' guanine dimer, the catalytic properties of SerHis were observed down to 8 μM (**Fig.2.16.**). This proves that although the activity of SerHis is small in comparison to modern day enzymes it can still be utilized in fairly low concentrations.

In the course of SerHis catalyzed oligomerization of ImpG, the formation of 3'-5' dimer and trimer were determined by coinjection with standards. Longer oligomers were not identified so precisely, however, within the chromatograms of the reactions the presence of various peaks with guanine UV-spectra beyond the trimer in the anion-exchange HPLC indicated that longer oligomers are also being formed. Thus SerHis catalysis is not limited to very short oligomers.

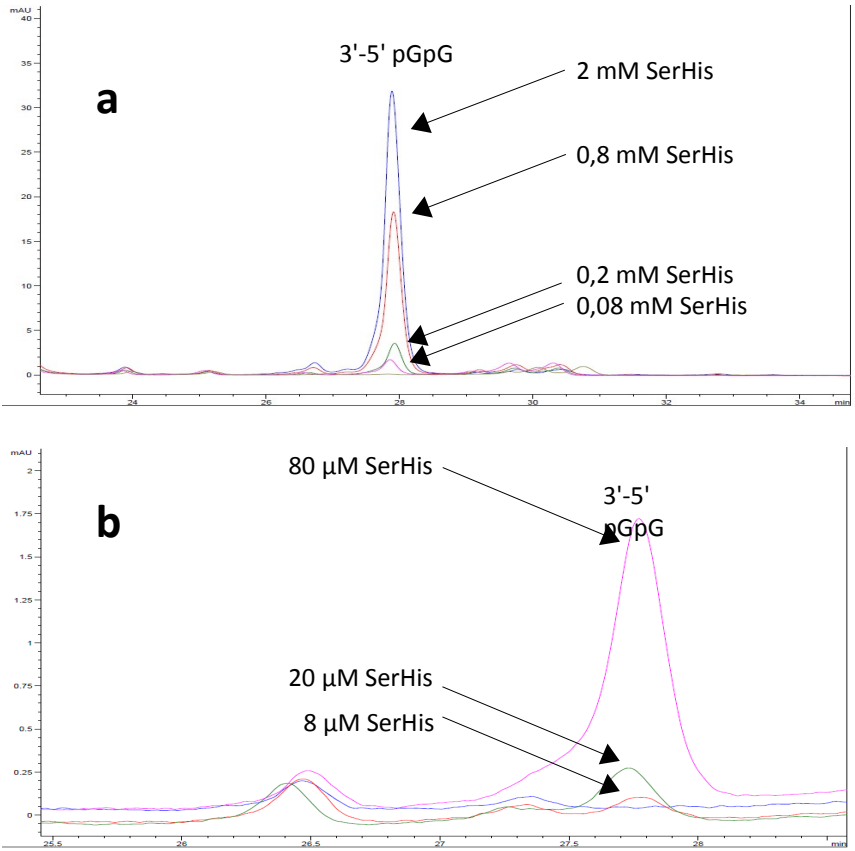


Figure 2.16. The formation 3'-5' guanine dimer with different concentrations of SerHis under standard conditions after 10 days and 25mM ImpG as starting substrate. **a** : 0,08-2 mM SerHis ; **b** : 8-80µM SerHis, the blue line represents the control chromatogram with no SerHis.

2.3. Investigation into the mechanism of SerHis catalysis

Oligomers of guanine summed up to only to 10% of the initial amount of ImpG. The major product of SerHis catalysis is GMP, around half of ImpG is transformed into GMP; GMP is a result of the hydrolysis of ImpG. SerHis catalysis is a balance between condensation and hydrolysis, therefore, almost 20% of the initial ImpG is transformed into a pyrophosphate dimer (**Fig.2.17.**). Pyrophosphate formation is not catalyzed and occurs spontaneously in a mixture of ImpG and GMP (Monnard *et al.* 2003).

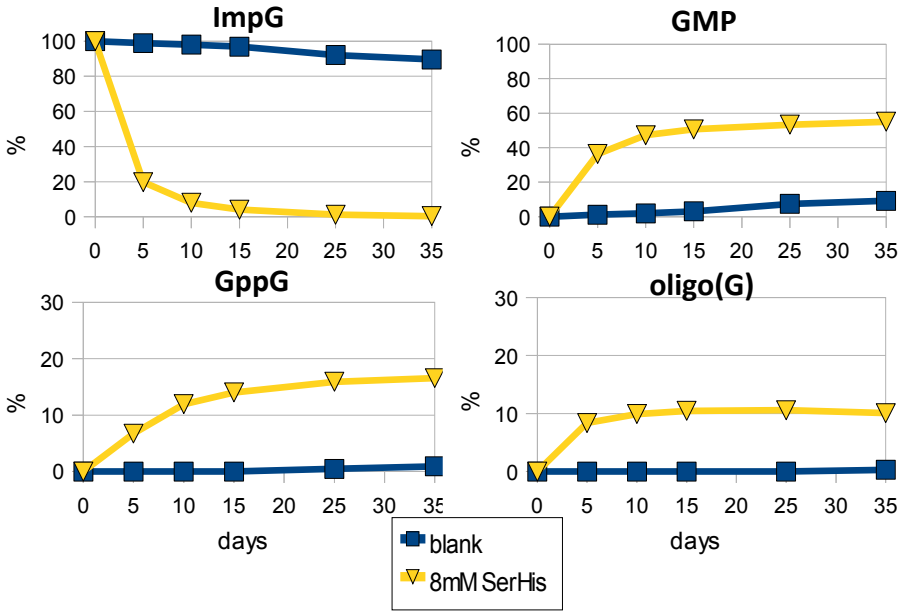


Figure 2.17. The formation of 1-SH-GMP and 2-SH-GMP compounds in the standard reaction mixture.

The molecular species presented in **Figure 2.14.** are the usual product of an ImpG self-condensation., however, they did not account for one hundred percent of initial ImpG. During the course of the incubation of reaction mixtures, the formation of two additional compounds in the monomer range of chromatograms was observed. Those compounds were

not present at the beginning of the reaction and were only formed in the samples containing SerHis. They had guanine UV spectra and monomeric range of retention in the anion-exchange HPLC. It can be reasonably assumed that those compounds were likely conjugates of SerHis and GMP. Therefore they were named 1-SH-GMP and 2-SH-GMP in order of their appearance on the chromatograms (**Fig.2.18.**).

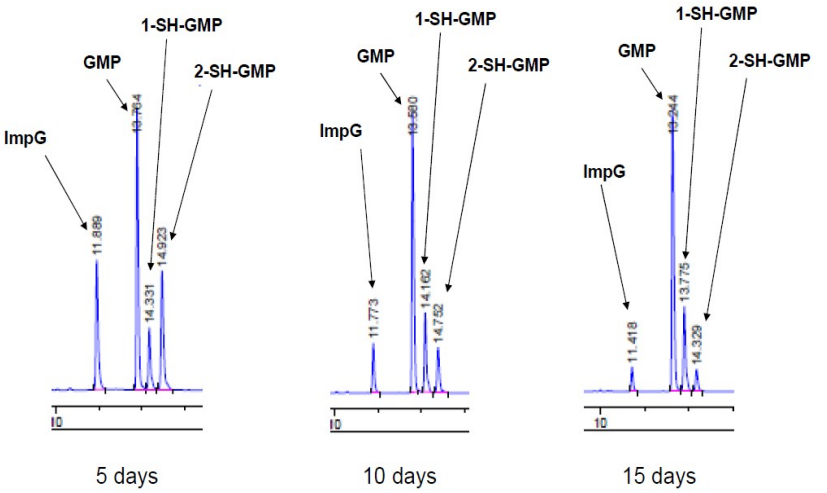


Figure 2.18. The formation of 1-SH-GMP and 2-SH-GMP compounds in the standard reaction mixture.

As can be observed in **Fig.2.18.** an interesting feature of 1-SH-GMP and 2-SH-GMP was that their relative concentrations changed over time, which indicates that they might be metastable compounds. Thus, their identification may shed more light on the mechanism by which SerHis catalyzes the formation of phosphodiester bonds.

Fig.2.19. and **Fig.2.20.** illustrate the relative change in the formation of both compounds. Compound 2-SH-GMP forms at the beginning of reaction and then slowly decreased, completely disappearing after around 30 days. Compound 1-SH-GMP formed slower, surpassing 2-SH-GMP around the 9th day and reaching plateau around 25th day. At the end of reaction 1-SH-GMP accounted for almost 20% of all guanine compounds in the system.

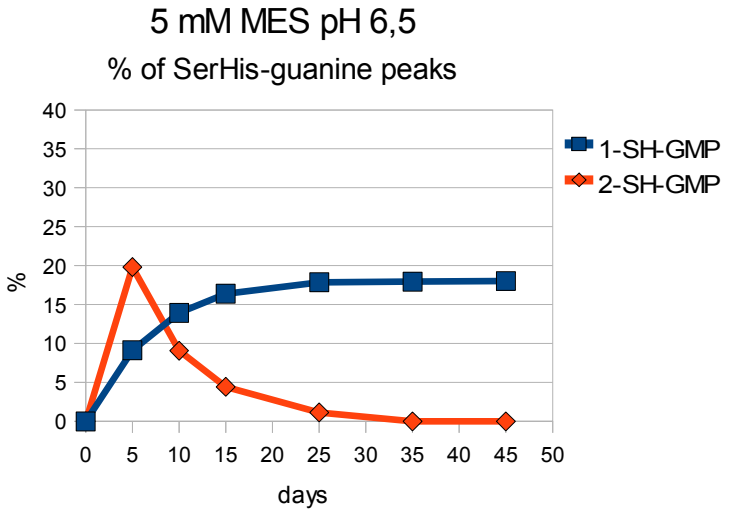


Figure 2.19. The formation of 1-SH-GMP and 2-SH-GMP compounds in the standard reaction mixture over 45 days.

The kinetics of 1-SH-GMP and 2-SH-GMP changed with pH (**Fig.2.17.**). The higher the pH, the greater the initial formation of 2-SH-GMP and the greater the final concentration of 1-SH-GMP. Also the equilibrium point between both compounds came later at higher pH.

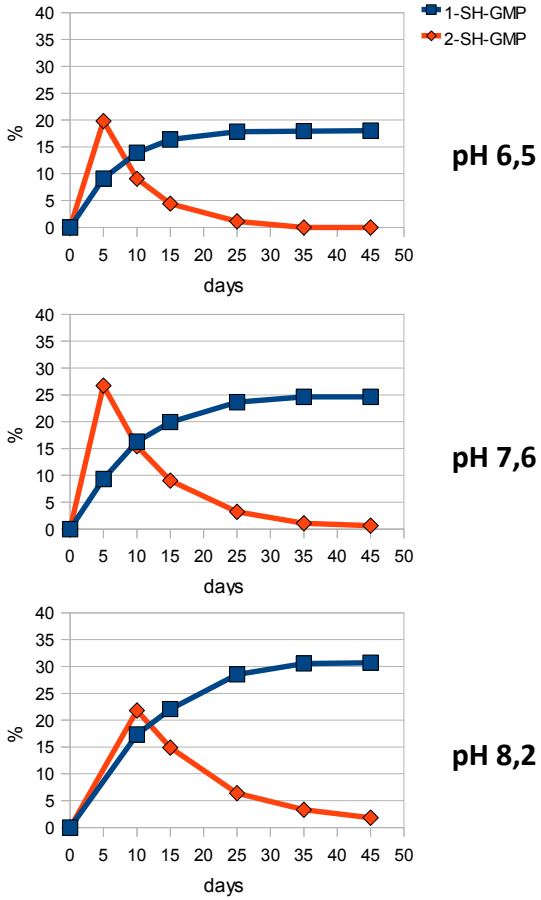


Figure 2.20. The formation of 1-SH-GMP and 2-SH-GMP compounds in different pH environments over 45 days.

More precise monitoring of the reaction during the initial period of incubation showed that the formation of 2-SH-GMP was very rapid (**Fig.2.21.**). Within the first hour almost 40% of ImpG was transformed into 2-SH-GMP, reaching a maximum yield of 44% after 8 hours. Afterwards, the amount of 2-SH-GMP slowly decreased.

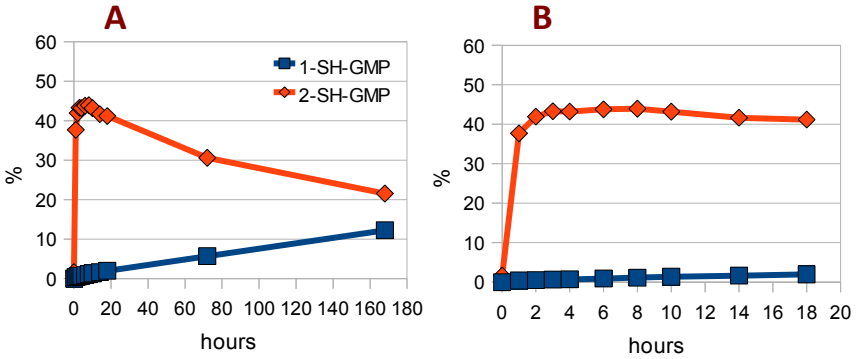


Figure 2.21. The formation of 1-SH-GMP and 2-SH-GMP compounds in the standard reaction mixture. **A:** reaction yields over a time period of 170 h, **B:** magnification of the first 20.

Both compounds 1-SH-GMP and 2-SH-GMP were isolated after separation on reverse phase C-18 HPLC column. The mass of the compounds was examined by MALDI mass spectrometry.

The molecular mass of SerHis is 242,2 g/mol and the molecular mass of GMP is 363,2 g/mol, therefore the mass of the SerHis/GMP conjugate should be 242,2+363,2-18,0 (water)=587,4 g/mol. The mass of both compounds was 587,3 g/mol (**Fig.2.22.**) thus proving that they were indeed covalent conjugates of SerHis and GMP.

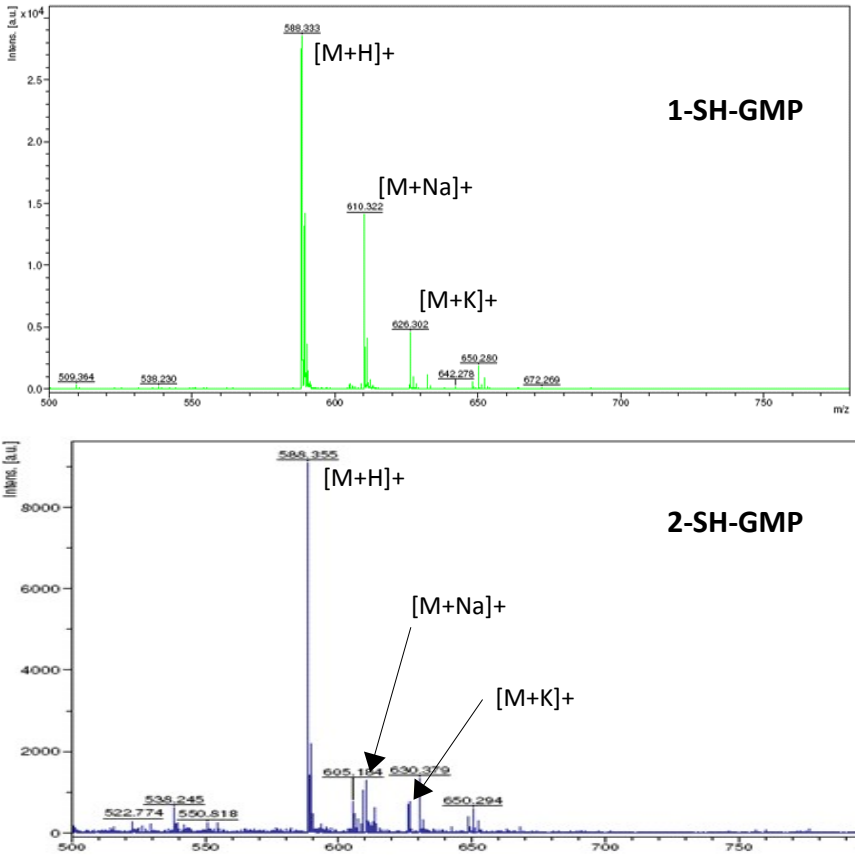


Figure 2.22. MALDI mass spectrometry of 1-SH-GMP and 2-SH-GMP. Both compounds have a mass of 587,3 g/mol.

From the kinetics of 1-SH-GMP and 2-SH-GMP we can postulate that 2-SH-GMP might be a reaction intermediate that is transformed into GMP and oligomers and 1-SH-GMP might be a kind of dead-end side reaction. In order to test this hypothesis, both compounds were isolated and used as the sole reactants in a standard reaction environment. During several days of incubation, 1-SH-GMP remained stable, whereas 2-SH-GMP was slowly consumed, producing GMP, 1-SH-GMP, pyrophosphate and the 3'-5' guanine dimer (**Fig.2.23.**).

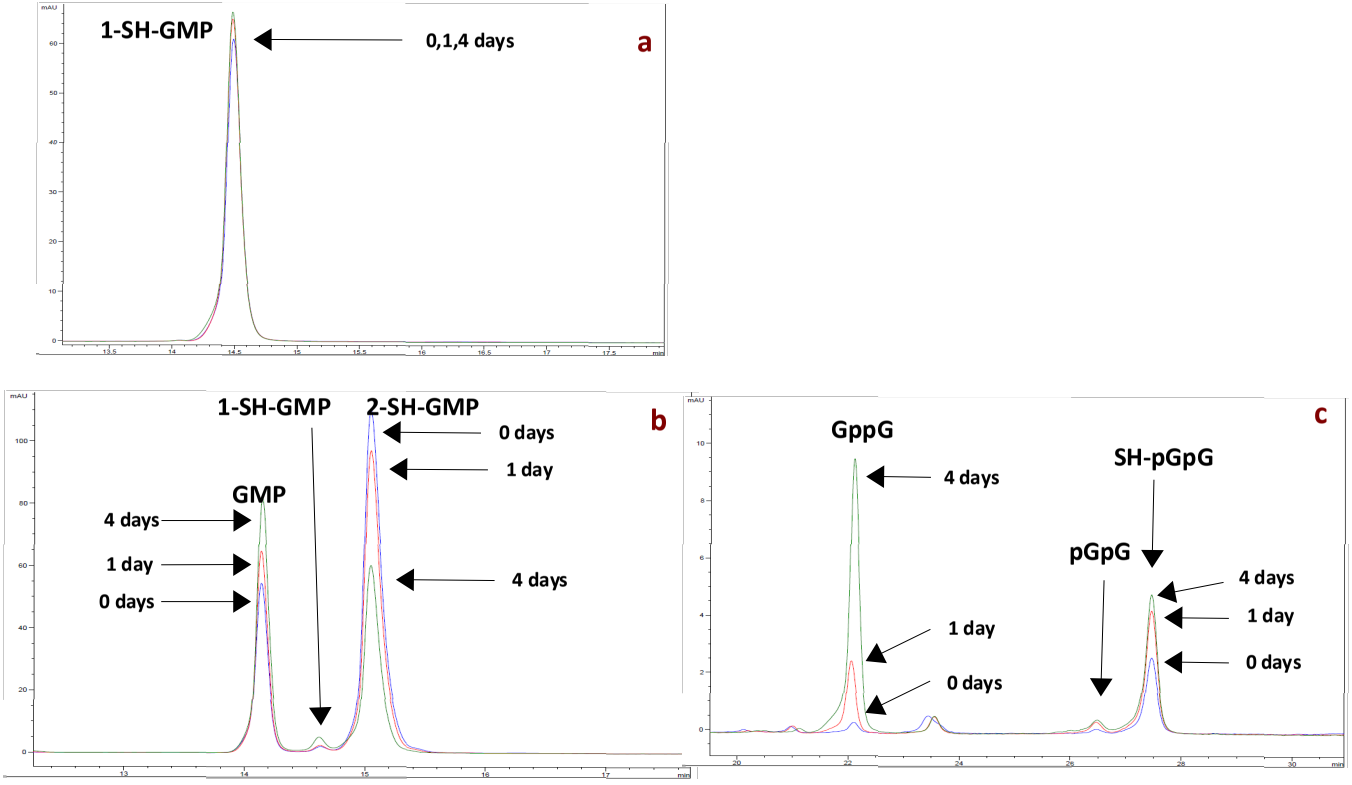


Figure 2.23. Chromatograms of 1-SH-GMP (a) and 2-SH-GMP (b,c) used as substrates for polymerization. **a** : 1-SH-GMP remained stable; **b** : 2-SH-GMP was consumed producing GMP and trace amounts of 1-SH-GMP; **c** : The conversion of 2-SH-GMP also resulted in the production of pyrophosphate and pGpG dimer.

This experiment proved that 2-SH-GMP is an intermediate leading to oligomers whereas 1-SH-GMP is an inactive side product. In comparison to the standard reaction, the amount of 1-SH-GMP formed is much smaller, suggesting that 1-SH-GMP does not rather form from 2-SH-GMP through intramolecular rearrangement but more likely is a product of independent attack of SerHis upon activated nucleotides, be it ImpG or 2-SH-GMP.

The remaining question is whether after decomposition of 2-SH-GMP the dipeptides SerHis is re-formed and able to participate in a new reaction, as a multiple turn-over true catalyst should. This issue is addressed in **Figure 2.24.** After almost complete decomposition of isolated 2-SH-GMP fresh ImpG was added and re-formation of 2-SH-GMP was observed thus proving the multiple turn-over capability of SerHis.

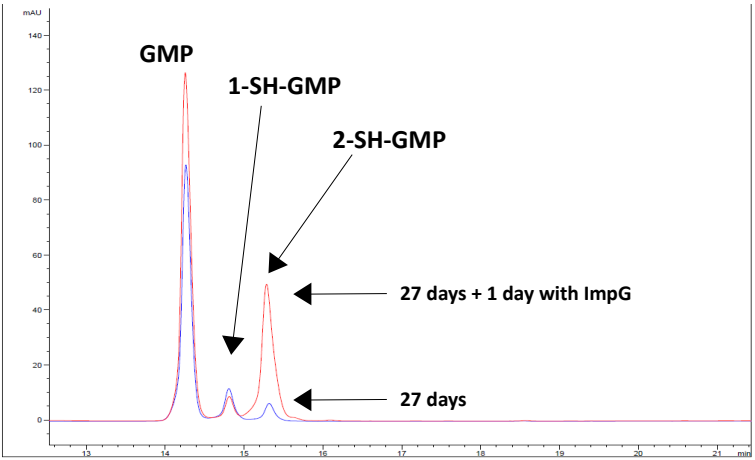


Figure 2.24. Chromatograms of isolated 2-SH-GMP almost completely transformed into other compounds after 27 days of incubation (blue line) and the same reaction one day after addition of fresh ImpG (red line).

There are three general ways in which the SerHis dipeptide can be joined to GMP (**Fig.2.25.**). The first is by formation of an ImpG analog in which the imidazole ring of ImpG is replaced by the imidazole ring of SerHis (**Fig.2.25. A**). There are presences in the literature of transamination on imidazole derivative activated ribonucleotides (Kanavarioti *et al.* 1995).

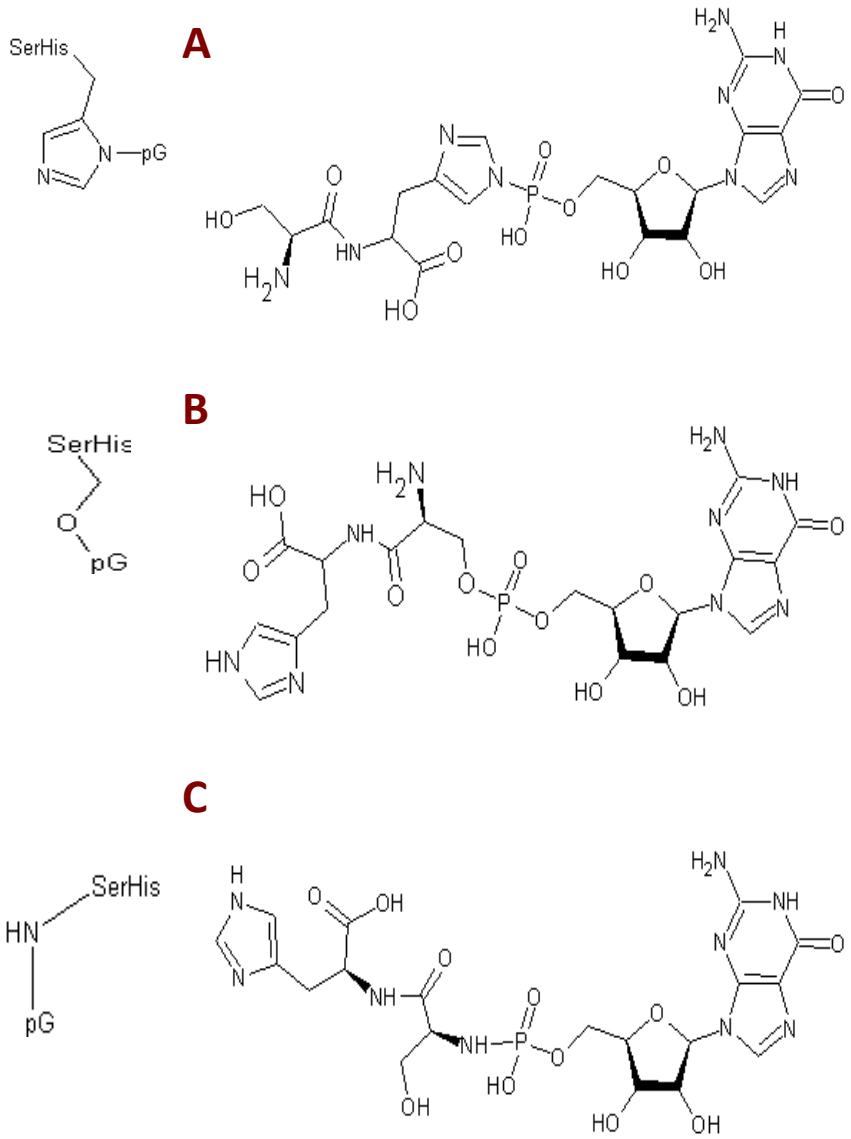


Figure 2.25 Three candidate structures for 1-SH-GMP and 2-SH-GMP. **A** : ImpG analog ; **B** : serine protease analog ; **C** : N-end joining.

The second option is an analog of an acyl-enzyme complex that is formed by serine proteases during cleavage of peptide bonds (**Fig.1.11.**). The linkage between GMP and SerHis goes through the hydroxyl oxygen of serine (**Fig.2.25.B**). The third possibility of linkage is through the amino end of peptide (**Fig.2.25.C**); primary amines are good nucleophiles therefore can attack phosphate group of ImpG and replace the imidazole.

In order to shed some light on the structural nature of 1-SH-GMP and 2-SH-GMP, a series of experiments was performed. SerHis was replaced by serine and its C- and/or N- blocked analogues (**Fig.2.27.**). The formation of new compound, assumed to be a covalent bond between GMP and serine, was observed in the case of unblocked serine and C-blocked derivative (H-Ser-OEt). However, a new compound was not observed in the case of both N- and C-blocked serine derivative (Ac-Ser-OMe). We can conclude that one of our compounds, namely 1-SH-GMP or 2-SH-GMP, is probably joined to GMP by its N-end (**Fig.2.25.C**). We can also exclude the fourth possibility (not presented in **Fig.2.25.**) of a C-end joining to GMP.

The resulting phospho-amide bond (**Fig.2.25.C**) should be relatively stable and its formation not only with SerHis but also with other compounds with free primary amine points to 1-SH-GMP (a stable side product) as a most likely candidate for this structure.

To further elucidate the mechanism of SerHis catalysis a standard reaction with a SerHis analog was prepared. SerHis(Met) is an analog of SerHis, with a methyl group in place of the hydrogen on the imidazole ring (**Fig.2.26.**). SerHis(Met) not only does not yield any oligomers but also forms only traces of the GMP bounded compound (**Fig.2.28.**). This points to the ImpG analog (**Fig.2.25. A**) as the most likely structure of 2-SH-GMP.

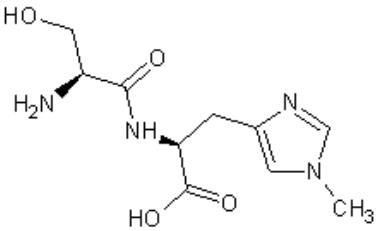


Figure 2.26. The structures SerHis(Met).

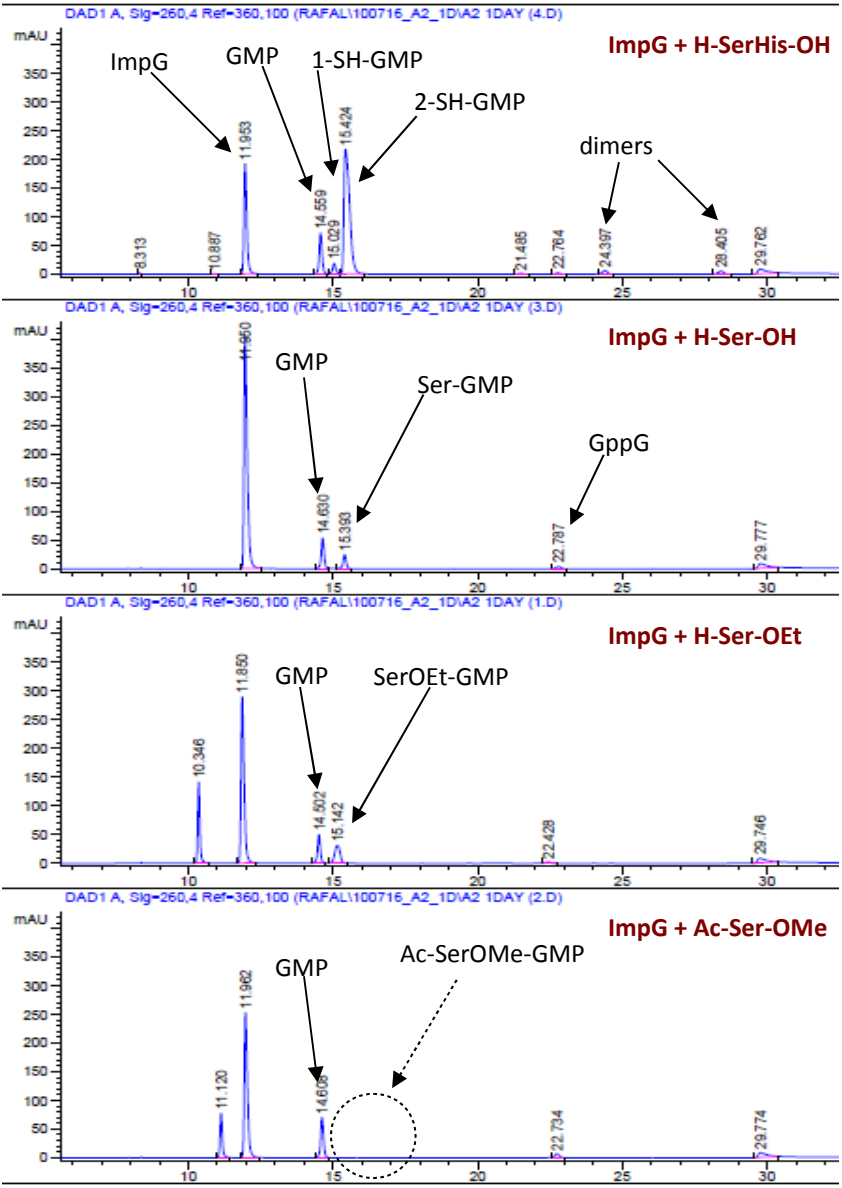


Figure 2.27. Interaction of ImpG with various reagents. Standard conditions, t = 1 day. Ac-Ser-OMe does not react with ImpG, contrary to H-Ser-OEt, H-Ser-OH and H-SerHis-OH.

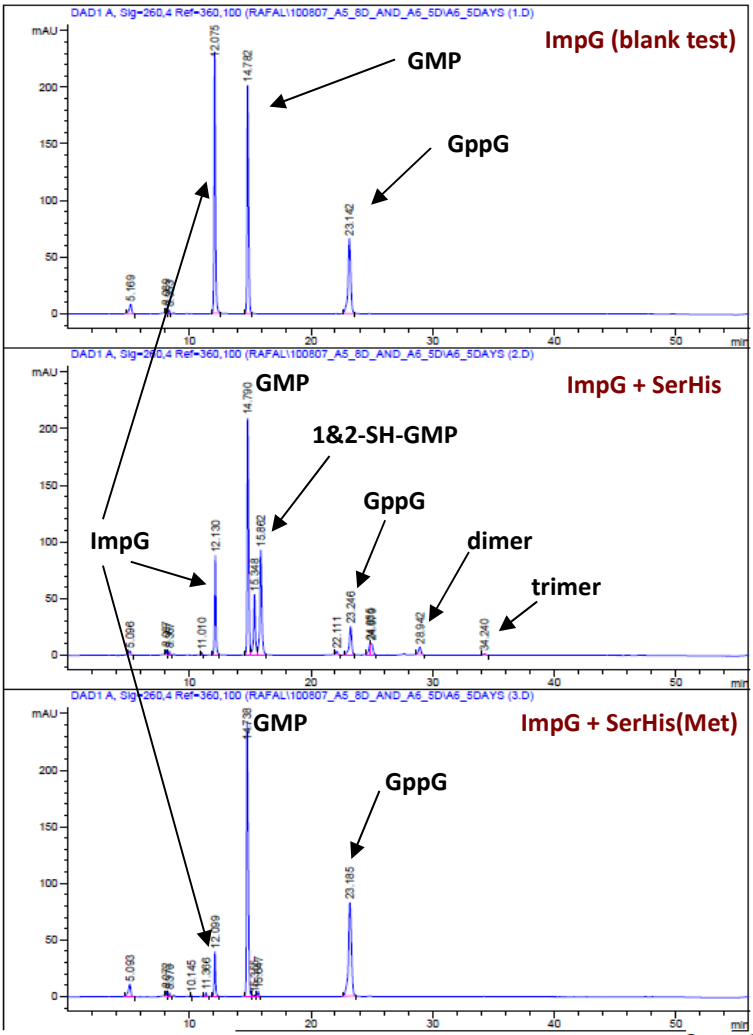


Figure 2.28. Interaction of ImpG with SerHis and SerHis(Met). Standard conditions, t = 5 day. SerHis(Met) does not yield any oligomers.

The titration curves highlight an interesting difference between SerHis and HisSer (**Fig.2.29.**). Both dipeptides have their three pK_a's close to the histidine values of 1,7 ; 6,0 and 9,1. However, the slope of deprotonation of the carboxyl group was much less steeper for SerHis than

it was for both HisSer and His. This might indicate that the C-end carboxylic group of SerHis is involved in a hydrogen bond, possibly with the imidazole group of the histidine residue.

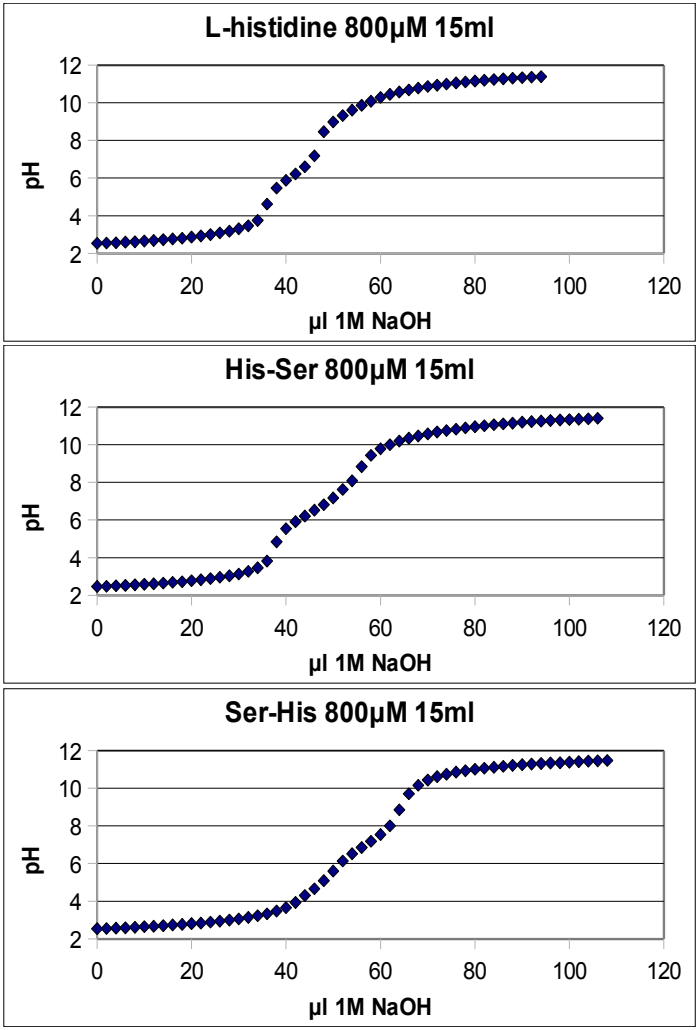


Figure 2.29. Titration curves for histidine, HisSer and SerHis.

3. Discussion

The present work has shown that the SerHis dipeptide can catalyze formation of nucleic acids from activated monomers. Oligomerization was achieved using different activating leaving groups in mononucleotides (**Fig.2.11.**). The catalytic effect was observed within broad pH values (5,0-8,2) and with three different buffers (**Fig.2.8.** and **Fig.2.9.**); these results show that the catalyst operates on general principles. SerHis catalyzed the formation of nucleic acid dimers with all four RNA nucleotides (**Fig.2.12.**). In addition to ribonucleotides, deoxyribonucleotides were also shown to be oligomerized by catalytic activity of SerHis (**Fig.2.13.**). Due to the lower nucleophilicity of the 3' hydroxyl group in comparison to the 2' hydroxyl group, non-enzymatic DNA oligomerization was not achieved previously in the eutectic phase in water/ice using metal-ion catalysts.

Although throughout most of the experiments shown in this thesis an equimolar concentration of catalyst and substrate was used (5 mM), a catalytic activity of SerHis was observed down to 8 μM concentration (**Fig.2.16.**). Moreover, SerHis proved to be a multiple turnover catalyst (**Fig.2.23.**); these properties point to SerHis as an ideal prebiotic analog of enzymes.

Although yields were relatively low, this is not detrimental in prebiotic scenarios as presumably, on the prebiotic earth, these reactions had thousands to millions of years to accumulate their products (Orgel 1998).

The efficiency of the SerHis catalysis is not high, around half of the activated monomer is transformed to its not activated form (that is GMP from the standard employed substrate ImpG). This is because after formation of the intermediate 2-SH-GMP, it can be attacked by different competing nucleophiles, which then release SerHis back into the environment. If it is attacked by water molecules the reactions results in deactivated GMP, if the attack is made by 2' or 3' hydroxyl groups of nucleotide's sugar it results in nucleic acid oligomers. The other alternative nucleophiles present in the system are: the phosphate group of the mononucleotide resulting in the formation of pyrophosphate dimer; the amino group from N-end of the peptide can also perform the attack resulting in the side product 1-SH-GMP. All those pathways are presented in **Figure 3.1.**

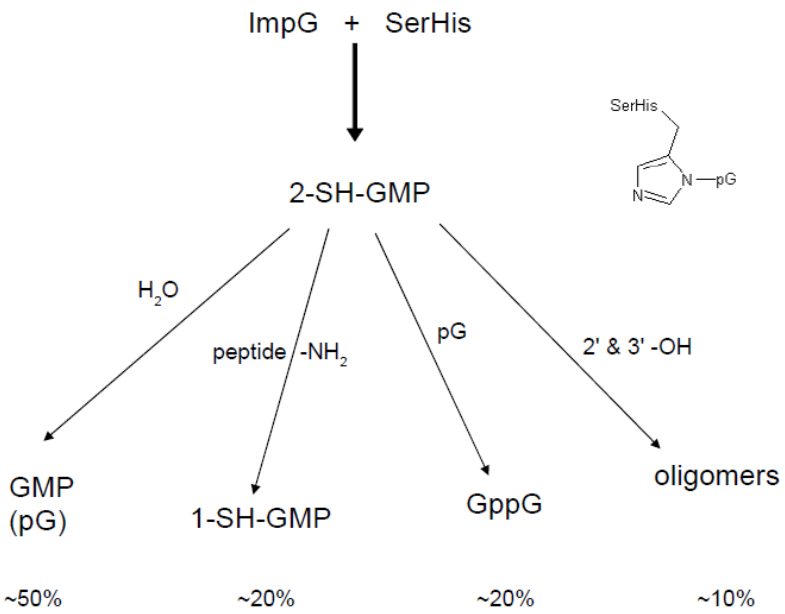


Figure 3.1. Products of SerHis interaction with activated mononucleotide substrate. The proposed nucleophiles attacking 2-SH-GMP are displayed next to the arrows. Oligomers constitute only around 10% of all products.

As with enzymes, the activity of SerHis results from the specific arrangement of different functional groups. Despite possessing all the functional groups of SerHis, the dipeptide HisSer and the modified dipeptide SerHis(Met) did not exhibit any catalytic activity.. Adding to this our knowledge on the mechanism of serine proteases (**Fig.1.11.**) we can safely speculate that the properties of SerHis result from the formation of one or more intramolecular hydrogen bonds. In most of the known serine proteases one hydrogen bond is formed between the rests of histidine and serine and a second hydrogen bond forms between the rests of histidine and aspartic acid. In our system we do not have the carboxyl group of aspartic acid. Serine proteases that contain only the Ser/His diad without the Asp residue are present in nature (Ekici *et al.* 2008), although their precise mechanism is not fully understood. Another explanation of SerHis activity could be that the carboxyl group of the peptides' C-end plays the same role

as aspartate residue in serine proteases. This would explain why HisSer does not possess the catalytic function of SerHis. In both cases, serine and histidine residues are in the same proximity and should be, in principle, equally prone to forming hydrogen bonds. However in the case of HisSer, the histidine residue is further away from the C-end carboxyl and therefore less likely to form hydrogen bonds. Different shapes of titration curves of both dipeptides points toward this possibility (**Fig.2.29.**). Also, the methylation of histidine would prevent the formation of a hydrogen bond, thus resulting in lack of catalysis (**Fig.2.28.**).

3.1. Prebiotic validity of presented reactions

Amino acids are easily formed in prebiotic conditions and many plausible scenarios has been proposed for their prebiotic condensation to peptides (chapter 1.2.). Serine is one of the most commonly observed amino acid products in prebiotic synthesis, while histidine is not encountered in most of the experiments and therefore is usually not considered a prebiotic amino acid. Nevertheless, a somewhat more indirect reaction for a prebiotically plausible production of histidine has been proposed by Shen (Shen *et al.* 1990a). Therefore, a peptide, especially as simple as dipeptide, composed of those two amino acids can rationally be considered as a plausible under prebiotic conditions. The synthesis of SerHis could have occurred following various synthetic pathways, for example in Salt Induced Peptide Formation (SIPF).

The prebiotic validity of frozen systems has been thoroughly explored (Trinks *et al.* 2005, Attwater *et al.* 2010). On contemporary Earth, substantial amounts of ice occur at the poles and at higher altitudes. Since the early Sun was less active than it is today, at least the same amount of ice is predicted to have been present on early Earth, and in the extreme case proposed by Bada *et al.* (1994), even the whole Earth would, at one point, have been entirely covered by an ice-sheet, similar to the presently existing ice covering Jovian moons Europa and Callisto (Schowman and Malhotra 1999), which posses a liquid water ocean beneath their icy surfaces. Even if this extreme proposition is not correct, substantial amounts of ice were very likely to have been present on early Earth, thus, the formation of eutectic water/ice environments similar to the one used in this work is entirely plausible.

Much more problematic in regards to the plausibility of SerHis acting as an enzymatic catalyst for nucleotide polymerization as proposed in this work is the availability of activated mononucleotides. Since the formation of phosphoimidazolides of nucleotides can be achieved relatively easily from imidazole and nucleotides, it was once claimed that they might be prebiotic compounds (Lohrmann 1977). However, presently their prebiotic formation is not regarded to be likely (Orgel 2004). As it was pointed in chapter 1.2.2. even the synthesis of simple unactivated nucleotides is not achievable under prebiotic conditions, with notable exception of very recent work of Sutherland's group (Powner *et al.* 2009). The subsequent activation of nucleotides is next and until now not solved missing link.

The lack of a prebiotic pathway to activated nucleotides is a generic weakness of non-enzymatic nucleic acid polymerization and the RNA-world in general (Lazcano and Miller 1996, Orgel 2004). Instead, the non-enzymatic polymerization of imidazole activated mononucleotides is the best currently available approximation of non-enzymatic nucleic acid polymerization that is presumed to have taken place at the dawn of the origin-of-life. This system allows us to study the kinetics and other properties of non-enzymatic polymerization, which differs in many features from biological enzymatic replication.

In the present study, imidazolides were also used as the best available approximation to that assumed by the RNA-world hypothesis as non-enzymatic replication of RNA material before the emergence of cells and their enzymatic apparatus.

3.2. Importance for the origin-of-life research

The fact that small small peptides can catalyze the formation of RNA is of tremendous importance for the field of prebiotic chemistry. Short peptides can reasonably be assumed to be prebiotically available compounds. On the contrary, nucleic acids do not form in prebiotic environments. The discovery that small peptides can have catalytic properties brought a new perspective to prebiotic chemistry. It was previously shown that the hydrolytic properties of SerHis can be reversed for peptide bonding (Gorlero *et al.* 2009) and now the same was shown for the phosphodiester bond.

The bio-organic chemistry approach in which unique properties of some products of prebiotic synthesis are used to create new compounds,

otherwise unachievable, has proven to be successful. This brings a new conceptual approach to the study of the origin-of-life. In contrast to prebiotic chemistry, this new approach is what might be called “prebiotic ecology”, since in this new approach we study the interactions of various chemical species in water environments in order to obtain the emergence of new properties which bring the system closer to minimal life.

The RNA world hypothesis postulates that modern life, which is based on DNA-RNA-proteins-DNA cycles, was preceded by a simpler system in which RNA played both the role of genotype (modern DNA) and phenotype (modern proteins) (Gilbert 1986). Such system could not, however, emerge directly in the prebiotic soup because RNA components are either not forming in prebiotic conditions or are not stable in such environments (Larralde *et al.* 1995). The ability of peptides to make RNA might be a way to bridge prebiotic chemistry and the RNA world (**Fig.3.2.**), thus bringing us closer to the understanding of origin-of-life.

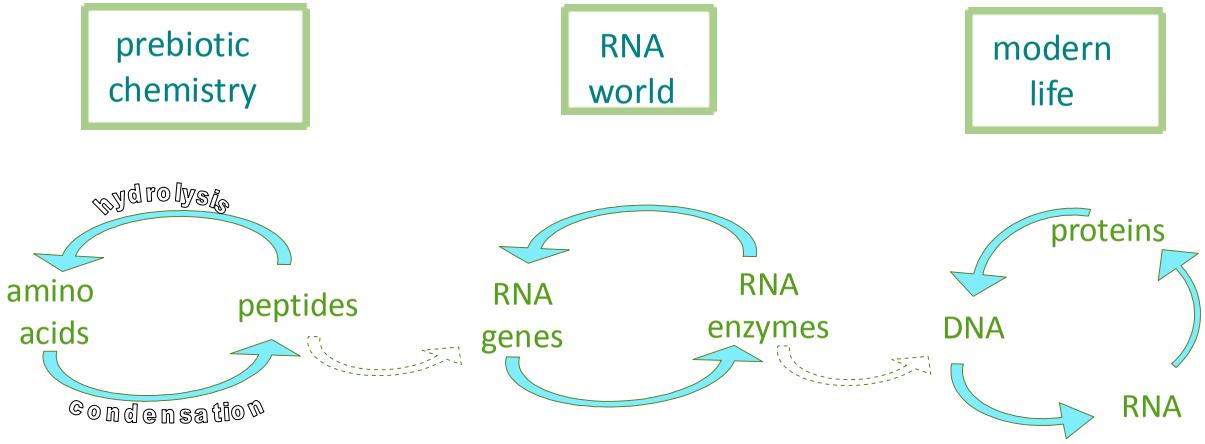


Figure 3.2. Proposed transitions from prebiotic chemistry to RNA world and from RNA world to modern life.

As we know, enzymes are very specific, highly efficient catalysts. We also know that they are products of the cumulative process of evolution over a time span of millions of years; their level efficiency and precision is a result of highly specific arrangements of amino acids' functional groups, such a specific arrangement is only possible when a polypeptide is long enough to fold in a way which brings together distant amino acids. These three notions lead to the conclusion that the ancestors of modern enzymes were less specific, less efficient and significantly shorter. The question that arises is: what was the minimum length of a functional peptide from which the evolution of modern enzymes could have begun? As it turns out, there is no minimal length, because even the smallest peptide – dipeptides - can posses catalytic activity. This is the case for the biggest enzyme family – serine hydrolases. The same might be true for other enzyme families.

Another interesting application of small peptides as prebiotic catalysts might be the etiology of specific macromolecular sequences. This problem is often neglected in origin-of-life research (Luisi 2007), nevertheless, it is important. How can we obtain a large number of, for example, an autocatalytic rybozyme? Till now only one experimental work has approached this issue (Chessari *et al.* 2006). The general problem lies in a generally random polymerization if one employes a non-enzymatic polymerization method. In such cases, acquiring more than one copy of a heteropolymer such as an RNA of the length, let us say thirty, is practically impossible. The best solution to this problem would be if a prebiotic catalyst could possess certain regioselectivity that would bias the polymerization toward certain sequences or a group of sequences. A peptide with all its functional groups pointing in various directions would likely posses some regioselectivity, thus opening the way for less random prebiotic polymerization.

Last but definitely not least, a new avenue opened when the hydrolytic peptide catalyzing condensation reactions were considered from the prospective of autocatalytic sets. It was hypothesized by Kauffman (1986) that, in a specific population of peptides of which some are hydrolytic and some are able to synthesize peptide bonds, the emergence of constantly increasing complexity would appear. This theory, supported only by calculations and considered to be an alternative to the RNA world, could not, however, be tested experimentally. However, with the discovery of the hydrolytic properties of SerHis and this dipeptide's ability to synthesize various chemical bonds, an experimental possibility for Kauffman's hypothesis can finally be conceived.

4. Conclusions and perspectives

In this thesis it was demonstrated that small peptide can catalyze the polymerization of nucleic acids in prebiotically compatible environments. This was achieved by employing the hydrolytic properties of the SerHis and designing experimental conditions in which this catalytic property was reversed, thus obtaining the condensation of mononucleotides into oligomeric products.

Polymerization of nucleic acids catalyzed by SerHis was observed in relatively broad conditions and with different substrates, thus establishing the general principle on which the activity of SerHis operates; SerHis properties depend on the structural arrangement of its functional groups, as the isomeric dipeptide HisSer did not exhibit any noticeable catalysis. In concentrations as low as 8 μM – a thousand time smaller than the substrate – the activity of SerHis was still detectable. SerHis was also proven to be a multiple turnover catalyst.

In order to have a better view of the processes here described some additional characterizations should be performed. A precise structural identification of the active intermediate compound, named in this work 2-SH-GMP should be performed. NMR techniques should be employed, however, after preparation and subsequent isolation a larger amount of material that is needed for successful NMR measurements. The maximum length of synthesized oligomers should also be determined, desirably by radioactive ATP labeling and polyacrylamide gel resolution. The same technique coupled with enzymatic digestion can be also used to determine the ratio between unnatural 2'-5' bonds and natural 3'-5' bonds in synthesized RNA .

Peptide catalyzed synthesis of RNA is of great importance for the RNA world hypothesis. Since RNA is not a product of known prebiotic pathways the reactions described in this thesis show a bridge between prebiotic chemistry and the RNA world and thanks to this we can, at least theoretically, envision smooth transitions from prebiotic chemistry through RNA world to modern life biochemistry.

Catalytic properties of small peptides are also of great importance for understanding the origin of enzymes and their role in the origin-of-life. Serine hydrolases, the biggest enzyme family throughout all domains of life,

seems to have a functional equivalent in a peptide as small as that comprised of only two amino acids. This shows that selection toward highly efficient catalytic peptides, which eventually resulted in present day enzymes, could have started at a very early stage of chemical evolution.

5. References

Attwater J, Wochner A, Pinheiro VB, Coulson A, Holliger P. (2010) Ice as a protocellular medium for RNA replication. *Nature Commun.* **1:** doi:10.1038/ncomms1076.

Bachem. (2009) Analytical Data Sheet. Lot number 0501749, H-Ser-His-OH acetate salt. https://www.bachem.com/ADS/0501749.pdf (accessed October 21, 2010).

Bada JL, Bigham C, Miller SL. (1994) Impact melting of frozen oceans on the early Earth: implications for the origin of life. *Proc Natl Acad Sci U S A.* **91:**1248-50.

Bada JL, Lazcano A. (2003) Perceptions of science. Prebiotic soup--revisiting the Miller experiment. *Science* **300(5620)**:745-6.

Barbier B, Brack A. (1987) Search for catalytic properties of simple polypeptides. *Orig Life Evol Biosph.* **17**:381-90.

Barbier B, Brack A. (1992) Conformation-Controlled Hydrolysis of Polyribonucleotides by Sequential Basic Polypeptides. *J. Am. Chem. Soc.* **114:** 3511-3515.

Bera PP., Nuevo M., Milam SN., Sandford SA., Lee TJ. (2010) Mechanism for the abiotic synthesis of uracil via UV-induced oxidation of pyrimidine in pure H2O ices under astrophysical conditions. J*ournal of Chemical Physics* **133**:104303-7.

Bernstein M. (2006) Prebiotic materials from on and off the early Earth. *Philos Trans R Soc Lond B Biol Sci.* **361:** 1689-700; discussion 1700-2.

Bordusa F. (2002) Proteases in organic synthesis. *Chem Rev.* **102:** 4817-68.

Bujdak J, Rode BM. (2004) On the mechanisms of oligopeptide reactions in solution and clay dispersion. *J Pept Sci.* **10:** 731-7.

Chen F, Yang D. (2007) Condensation of amino acids to form peptides in aqueous solution induced by the oxidation of sulfur(iv): an oxidative model

for prebiotic peptide formation. *Orig Life Evol Biosph.* **37:** 47-54.

Chen J, Wan R, Liu H, Cheng C, Zhao Y. (2001) Cleavage of BSA by a dipeptide seryl-histidine. *Letters in Peptide Science* **7:** 325–329.

Chessari S, Thomas RM, Polticelli F, Luisi PL. (2006) The Production of de novo Folded Proteins by a Stepwise Chain Elongation: A Model for Prebiotic Chemical Evolution of Macromolecular Sequences. *Chemistry & Biodiversity* **3:**1202-1210.

Cleaves HJ 2nd. (2010) The origin of the biologically coded amino acids. *J Theor Biol.* **263**:490-8.

Commeyras A, Collet H, Boiteau L, Taillades J, Vandenabeele-Trambouze O, Cottet H, Biron J-P, Plasson R, Mion L, Lagrille O. (2002) Prebiotic synthesis of sequential peptides on the Hadean beach by a molecular engine working with nitrogen oxides as energy sources. *Polymer International* **51:** 661.

Cronin JR, Pizzarello S. (1997) Enantiomeric excesses in meteoritic amino acids. *Science* **275:** 951-5.

Decker P, Schweer H, Pohlmann R. (1982) Identification of formose sugars, presumable prebiotic metabolites, using capillary gas chromatography/gas chromatography-mass spectrometry of *n*butoxime trifluoroacetates on OV-225. *J Chromatog* **244:**281–291.

Du H, Wang Y, Yang L, Luo W, Xia N, Zhao Y. (2002) Appraisal of green fluorescent protein as a model substrate for seryl-histidine dipeptide cleaving agent. *Letters in Peptide Science* **9:** 5-10.

Dziedzic P, Zou W, Hafren J, Cordova A. (2006) The small peptide-catalyzed direct asymmetric aldol reaction in water. *Org Biomol Chem.* **4:** 38-40.

Ekici OD, Paetzel M, Dalbey RE. (2008) Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci.* **17:**2023-37.

Ferris JP, Sanchez RA, Orgel LE. (1968) Studies in prebiotic synthesis. 3.

Synthesis of pyrimidines from cyanoacetylene and cyanate. *J Mol Biol* **33:**693–704.

Ferris JF, Joshi PC, Edelson EH, Lawless JG. (1978) HCN: a plausible source of purines, pyrimidines and amino acids on the primitive Earth. *J. Mol. Evol.* **11:**293–311.

Fleminger G, Yaron T, Eisenstein M, Bar-Nun A. (2005) The structure and synthetic capabilities of a catalytic peptide formed by substrate-directed mechanism – implications to prebiotic catalysis. *Origins of Life and Evolution of Biospheres* **35:** 369–382.

Fox SW, Harada KJ. (1960) The Thermal Copolymerization of Amino Acids Common to Protein. *J Am Chem Soc* **82:** 3745.

Friedmann N, Miller SL. (1969) Phenylalanine and tyrosine synthesis under primitive earth conditions. *Science* **166:**766–767.

Gesteland RF, Cech TR, Atkins JF. (1999) *The RNAWorld, 2*nd ed. Cold Springs Harbor, NY: Cold Springs Harbor Press.

Gilbert W. (1986) The RNA World. *Nature* **319:**618.

Gorlero M, Wieczorek R, Adamala K, Giorgi A, Schininà ME, Stano P, Luisi PL. (2009) Ser-His catalyses the formation of peptides and PNAs. *FEBS Lett.* **583:**153-6.

Greenstein JP, Winitz M. (1996) *Chemistry of the Amino Acids* Vol. 2 s.861. Wiley and Sons, New York.

Hill AR Jr, Orgel LE, Wu T. (1993) The limits of template-directed synthesis with nucleoside-5-phosphoro(2-methyl)imidazolides. *Orig Life Evol Biosph* **23:**285–290.

Huang W, Ferris JP. (2003) Synthesis of 35-40 mers of RNA oligomers from unblocked monomers.Asimple approach to theRNA world. *Chem Commun (Camb)* **12:**1458–1459.

Ingar AA, Luke RW, Hayter BR, Sutherland JD. (2003) Synthesis of cytidine ribonucleotides by stepwise assembly of the heterocycle on a sugar phosphate. *Chembiochem.* **4**:504-7.

Itoh T, Tomizawa J. (1980) Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *PNAS USA* **77**:2450-4.

Joyce GF. (1987) Nonenzymatic template-directed synthesis of informational macromolecules. *Cold Spring Harb Symp Quant Biol* **52:**41–51.

Joyce GF. (1989) RNA evolution and the origins of life. *Nature* **338**:217-24.

Kanavarioti A, Stronach MW, Ketner RJ, Hurley TB. (1995) Large steric effect in the substitution reaction of amines with phosphoimidazolide-activated nucleosides. *J Org Chem.* **60:**632-7.

Kanavarioti A, Lee LF, Gangopadhyay S. (1999) Relative reactivity of ribosyl 2'-OH vs. 3'-OH in concentrated aqueous solutions of phosphoimidazolide activated nucleotides. *Orig Life Evol Biosph.* **29:**473-87.

Kanavarioti A, Monnard PA, Deamer DW. (2001) Eutectic phases in ice facilitate nonenzymatic nucleic acid synthesis. *Astrobiology* **1:**271–281.

Kauffman SA. (1986) Autocatalytic set of proteins. *J. Theor. Biol.* **119:**1-24.

Kenner GW, Stedman RJ. (1952) Peptides. Part I. The Synthesis of Peptides through Anhydrides of Sulphuric Acid. *J Chem Soc* 2069–2076.

Kofoed J, Nielsen J, Reymond JL. (2003) Discovery of new peptide-based catalysts for the direct asymmetric aldol reaction. *Bioorg Med Chem Lett.* **13:** 2445-7.

Kraut J. (1977) Serine proteases: structure and mechanism of catalysis. *Annu Rev Biochem.* **46:** 331-58.

Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31:**147–157.

Lahav N, White D, Chang S. (1978) Peptide formation in the prebiotic era: thermal condensation of glycine in fluctuating clay environments. *Science* **201:** 67-69.

Larralde R, Robertson MP, Miller SL. (1995) Rates of decomposition of ribose and other sugars: implications for chemical evolution. *Proc Natl Acad Sci U S A.* **92:**8158-60.

Lazcano A, Miller SL. (1996) The origin and early evolution of life: prebiotic chemistry, the pre-RNA world, and time. *Cell* **85:**793-8.

Leman L, Orgel L, Ghadiri MR. (2004) Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **306:** 283-6.

Li Y, Zhao Y, Hatfield S, Wan R, Zhu Q, Li X, McMills M, Ma Y, Li J, Brown KL, He C, Liu F, Chen X. (2000) Dipeptide seryl-histidine and related oligopeptides cleave DNA, protein, and a carboxyl ester. *Bioorg Med Chem.* **8:** 2675-80.

List B, Lerner RA, Barbas CF 3rd. (2000) Proline-Catalyzed Direct Asymmetric Aldol Reactions. *J. Am. Chem. Soc.* **122:** 2395-2396.

Lohrmann, R. (1977) Formation of nucleoside 5-phosphoramidates under potentially prebiological conditions. *J Mol Evol* **10:**137–154.

Luisi PL. (1979) Why are enzymes macromolecules? *Naturwissenschaften* **66:**498-504.

Luisi PL. (2006) *The Emergence of Life: from Chemical Origin to Synthetic Biology.* ss. 56-57. Cambridge University Press.

Luisi PL. (2007) Question 3: The Problem of Macromolecular Sequences: The Forgotten Stumbling Block. *Orig Life Evol Biosph* **37:** 363-365.

Ma Y, Chen X, Sun M, Wan R, Zhu C, Li Y, Zhao Y. (2007) DNA cleavage function of seryl-histidine dipeptide and its application. *Amino Acides* **35:**251-6.

Marvin MC, Engelke DR. (2009) Broadening the mission of an RNA enzyme. *J Cell Biochem.* **108:**1244-51.

Martins Z, Botta O, Fogel ML, Sephton MA, Glavin DP, Watson JS, Dworkin JP, Schwartz AW, Ehrenfreund P (2008). Extraterrestrial nucleobases in the Murchison meteorite. *Earth and Planetary Science Letters* **270:**130–136.

Miller SL. (1953) A production of amino acides under possible primitive Eart conditions. *Science* **117:** 528.

Miller SL. (1993) The prebiotic synthesis of organic compounds on the Early Earth. In: Michael, H.E., Stephen, A.M. (Eds.), *Organic Geochemistry: Principles and Applications*. Plenum Press, New York.

Miyakawa S, Cleaves HJ, Miller SL. (2002) The cold origin of life: B. Implications based on pyrimidines and purines produced from frozen ammonium cyanide solutions. *Orig Life Evol Biosph.* **32:**209-18.

Mizuno T. Weiss AH. (1974) Synthesis and utilization of formose sugars. In: *Advances in Carbohydrate Chemistry and Biochemistry*, v. 29, pp. 173–227. Tipson, R.W., and Horton, D., Eds., New York London: Academic Press.

Monnard PA, Kanavarioti A, Deamer DW. (2003) Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. *J Am Chem Soc.* **125:**13734-40.

Monnard PA, Szostak JW. (2008) Metal-ion catalyzed polymerization in the eutectic phase in water-ice: a possible approach to template-directed RNA polymerization. *J Inorg Biochem.* **102:**1104-11.

Monnard PA, Ziock H. (2008) Eutectic phase in water-ice: a self-assembled environment conducive to metal-catalyzed non-enzymatic RNA polymerization. *Chem Biodivers.* **8:**1521-39.

Moravek J. (1967) Formation of oligonucleotides during heating of a mixture of uridine 2(3)-phosphate and uridine. *Tetrahedron Lett* **18:**1707–1710.

Moser RE, Claggett AR, Matthews CN. (1968) Peptide formation from aminomalonitrile (HCN trimer). *Tetrahedron Lett* **13:** 1605–1608.

Nelson KE, Levy M, Miller SL. (2000) Peptide nucleic acids rather than RNA may have been the first genetic molecule. *Proc Natl Acad Sci U S A.* **97:**3868-71.

Nelson KE, Robertson MP, Levy M, Miller SL. (2001) Concentration by evaporation and the prebiotic synthesis of cytosine. *Orig Life Evol Biosph* **31:**221–229.

O'Donoghue P, Luthey-Schulten Z. (2003) On the evolution of structure in aminoacyl-tRNA synthetases. *Microbiol Mol Biol Rev.* **67:**550-73.

Orgel LE. (1998) The origin of life--how long did it take? *Orig Life Evol Biosph.* **28:**91-6.

Orgel LE. (2004) Prebiotic chemistry and the origin of the RNA world. *Crit Rev Biochem Mol Biol.* **39:**99-123.

Oro J. (1961) Mechanism of synthesis of adenine from hydrogen cyanide under possible primitive earth conditions. *Nature* **191:**1193-4.

Pizzarello S, Huang Y, Fuller M, (2004) The carbon isotopic distribution of Murchison amino acids. *Geochim. Cosmochim. Acta* **68:**4963–4969.

Pizzarello S, Weber AL. (2004) Prebiotic amino acids as asymmetric catalysts. *Science* **303:** 1151.

Pizzarello S, Shock E. (2010) The organic composition of carbonaceous meteorites: the evolutionary story ahead of biochemistry. *Cold Spring Harb Perspect Biol.* **2:**a002105.

Plankensteiner K, Righi A, Rode BM. (2002) Glycine and diglycine as possible catalytic factors in the prebiotic evolution of peptides. *Orig Life Evol Biosph.* 32: 225-36.

Powner MW, Gerland B, Sutherland JD. (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459:**239-42.

Prabahar KJ, Ferris JP. (1997) Adenine derivatives as

phosphateactivating groups for the regioselective formation of 3,5-linked oligoadenylates on montmorillonite: possible phosphate-activating groups for the prebiotic synthesis ofRNA. *JAmChem Soc* **119:**4330–4337.

Rajamani S, Vlassov A, Benner S, Coombs A, Olasagasti F, Deamer D. (2008) Lipid-assisted synthesis of RNA-like polymers from mononucleotides. *Orig Life Evol Biosph.* **38:**57-74.

Robertson MP, Miller SL. (1995) An efficient prebiotic synthesis of cytosine and uracil. *Nature* **375:**772–774.

Rode BM. (1999) Peptides and the origin of life. *Peptides* **20:** 773–786.

Rode BM, Son HL, Suwannachot Y, Bujdak J. (1999) The combination of salt induced peptide formation reaction and clay catalysis: a way to higher peptides under primitive earth conditions. *Orig Life Evol Biosph.* **29:** 273-86.

Sawai H, Orgel LE. (1975) Oligonucleotide synthesis catalyzed by the Zn2+ ion. *J Am Chem Soc* **97:**3532-3533.

Schrum JP, Ricardo A, Krishnamurthy M, Blain JC, Szostak JW. (2009) Efficient and rapid template-directed nucleic acid copying using 2'-amino-2',3'-dideoxyribonucleoside-5'-phosphorimidazolide monomers. *J Am Chem Soc.* **131:**14560-70.

Schuster M, Aaviksaar A, Jakubke HD. (1990) Enzyme-catalyzed peptide synthesis in ice. *Tetrahedron* **46:**8083-8102.

Schwendinger MG, Rode BM. (1989) Possible role of copper and sodium chloride in prebiotic evolution of peptides. *Analytical Sciences* **5:**411-414.

Schwendinger MG, Rode BM. (1992) Investigations on the mechanism of the salt-induced peptide formation. *Orig Life Evol Biosph.* **22:** 349-59.

Shen C, Yang L, Miller SL, Oró J. (1990a) Prebiotic synthesis of histidine. *Journal of Molecular Evolution* **31:**167-174 .

Shen C, Lazcano A Oró J. (1990b) The enhancement activites of histidyl-histidine in some prebiotic reactions. *Journal of Molecular Evolution* **31:** 445-452.

Showman AP, Malhotra R. (1999). The Galilean Satellites. *Science* **286:**77–84.

Sigma-Aldrich (2010) PolyAmino Acids: Production Methods FAQ http://www.sigmaaldrich.com/life-science/metabolomics/learning-center/polyamino-acids-production-faq.html (accessed October 21, 2010).

Simon GM, Cravatt BF. (2010) Activity-based proteomics of enzyme superfamilies: serine hydrolases as a case study. *J Biol Chem.* **285:**11051-5.

Sun M, Chen J, Liu XH, Zhao YF. (2004) Molecular interaction modeling of Ser-His dipeptide and buffers. *Journal of Molecular Structure: THEOCHEM* **668:**47-49.

Szostak JW. (2009) Origins of life: Systems chemistry on early Earth. *Nature* **459:**171-2.

Taillades J, Beuzelin I, Garrel L, Tabacik V, Bied C, Commeyras A. (1998) N-carbamoyl-alpha-amino acids rather than free alpha-amino acids formation in the primitive hydrosphere: a novel proposal for the emergence of prebiotic peptides. *Orig Life Evol Biosph.* **28:** 61-77.

Taillades J, Collet H, Garrel L, Beuzelin I, Boiteau L, Choukroun H, Commeyras A. (1999) N-carbamoyl amino acid solid-gas nitrosation by NO/NOx: A new route to oligopeptides via alpha-amino acid N-carboxyanhydride. Prebiotic implications. *J Mol Evol.* **48:** 638-45.

Trinks H, Schröder W, Biebricher CK. (2005) Ice and the origin of life. *Orig Life Evol Biosph.* **35:**429-45.

Van Trump JE, Miller SL. (1972) Prebiotic synthesis of methionine. *Science* **178:**859–860.

Yamagata Y. (1999) Prebiotic formation of ADP and ATP from AMP, calcium phosphates and cyanate in aqueous solution. *Orig Life Evol Biosph* **29:**511–520.

Yanagawa H, Kobayashi K. (1992) An experimental approach to chemical evolution in submarine hydrothermal systems. *Orig Life Evol Biosph* **22:**

147-159.

Yanagawa H, Kojima K, Ito M, Handa N. (1990) Synthesis of polypeptides by microwave heating I. Formation of polypeptides during repeated hydration-dehydration cycles and their characterization. *J Mol Evol.* **31:** 180-6.

Wakamatsu, H., Yamada, Y., Saito, T., Kumashiro, I., and Takenishi, T. 1966. Synthesis of adenine by oligomerization of hydrogen cyanide. *J Org Chem* **31:**2035–2036.

Weber AL, Pizzarello S. (2006) The peptide-catalyzed stereospecific synthesis of tetroses: a possible model for prebiotic molecular evolution. *Proc Natl Acad Sci USA*. **103:** 12713-7.

White HB 3rd. (1976) Coenzymes as fossils of an earlier metabolic state. *J Mol Evol.* **7:**101-4.

White DH, Erickson JC. (1980) Catalysis of peptide bond formation by histidyl-histidine in a fluctuating clay environment. *Journal of Molecular Evolution* **16:** 279-290.

Woese CR (1967) *The genetic code*. Harper & Row, New York

Zou W, Ibrahem I, Dziedzic P, Sundén H, Córdova A. (2005) Small peptides as modular catalysts for the direct asymmetric aldol reaction: ancient peptides with aldolase enzyme activity. *Chem Commun (Camb)*. **39:** 4946-8.

Zubay, G. (1998) Studies on the lead-catalyzed synthesis of aldopentoses. *Orig Life Evol Biosph* **28:**13–26.