UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Chemistry

Stanislav Andres

Development of method for preliminary identification of cyclic dinucleotides in bacterial cultures

Master's Thesis
(30 ECTS in Applied Measurement Science)

Supervisors:
Prof. Dr. Jörg Hartig
Faculty of Science
Department of Chemistry
University of Konstanz
Associate professor Koit Herodes
Faculty of Science and Technology
Institute of Chemistry
University of Tartu

TARTU
2016

CONTENTS

LIST OF ABREVIATIONS

C18 GravitySB (C18 GSB) – liquid chromatography C18-based column GravitySB™

C18 Isis – liquid chromatography C18-based column Isis™

c-AMP-CMP (cACMP) – cyclic adenosine monophosphate cytosine monophosphate

c-AMP-GMP (cAGMP) – cyclic adenosine monophosphate guanosine monophosphate

c-AMP-UMP (cAUMP) – cyclic adenosine monophosphate uridine monophosphate

c-CMP-GMP (cCGMP) – cyclic cytosine monophosphate guanosine monophosphate

c-CMP-UMP (cCUMP) – cyclic cytosine monophosphate uridine monophosphate

c-di-AMP – cyclic diadenosine monophosphate

c-di-CMP – cyclic dicytosine monophosphate

c-di-GMP – cyclic diguanosine monophosphate

c-di-NMP (CDN) – cyclic dinucleotide monophosphate

c-di-UMP – cyclic diuridine monophosphate

c-GMP-UMP (cGUMP) – cyclic guanosine monophosphate uridine monophosphate

HPLC – high-performance liquid chromatography

LC-MS – liquid chromatography – mass spectrometry

LOD – limit of detection

NMR – nuclear magnetic resonance

OD_{600} – optical density at wavelength of 600 nm

PDE – phosphodiesterase

INTRODUCTION

Nucleotides, besides being monomers of nucleic acids, are widespread second messengers of all the cellular organisms. Among them, a new group – cyclic dinucleotides (CDN) – is recently getting high attention, as the processes, they take part in, are found to be highly important for organisms (Romling *et al.* 2013; Sun *et al.* 2013; Huynh and Woodward 2016). Currently, only three of these compounds were confirmed in living cells (Kalia *et al.* 2013). The methods for their identification are varying and do not allow to compare directly the amount of compounds between them, as well as, they are not optimized to work with several of the compounds (Spangler *et al.* 2010; Gao *et al.* 2015). Another important point, is that currently the field of the cyclic dinucleotides is rapidly expanding and the new proteins, predicted to utilize CDNs are discovered literary every day (Krasteva *et al.* 2012; Tschowri 2016; Hallberg *et al.* 2016). However, many of them do not show activity towards known compounds or their metabolites (Huynh and Woodward 2016). This suggests possible presence of different compounds of cyclic dinucleotides group in living cell.

Therefore, it is important to develop a method, that can reveal such compounds and be ubiquitous towards all known and proposed cyclic dinucleotide compounds. In addition, as without highly pure standard materials it is hard to set up high-precision quantification method, it is still possible to develop one for preliminary identification of such compounds in cells for later biological use.

Based on said, the aims of the study are the following:

- Set up the experimental method to analyze mixture of synthetic cyclic dinucleotide compounds,
- Develop sample preparation and bacteria growing procedure to keep possible compounds intact,
- Perform analysis of living cells to confirm the ability of the method to identify known compounds and the candidates for new ones.

1. LITERATURE OVERVIEW

1.1. Historical Perspective

Nucleic acids and their monomers – nucleotides – have been playing central role in all the living organisms throughout the whole history of life. From the simplest RNA-like oligomers and triphosphates to all the currently known and still unknown forms, these molecules were and still are guiding the development of all *biota*. Not only the genetic information storage is the feature of these compounds, but also variety of metabolic and signaling functions.

The first important step was discovery of ATP (adenosine-triphosphate) and its functions by Karl Lohmann 1929 and Fritz Albert Lipmann 1941 respectively (Florkin 1991). These discoveries gave the biological society the first good candidate for the energy source of the mammals (later, all the cellular forms of life). Later other triphosphonucleotides where shown to have the similar functions.

DNA primal structure, discovered by James Watson and Francis Crick in 1953, whose model-building efforts were guided by X-ray diffraction data, was a milestone in the whole biological approach and started the new branch of Life science – molecular biology (Lehmann 2003).

Next key point was the discovery of adenosine-containing compound that initiates cellular response to epinephrine and glucagon in liver homogenate. Using UV-spectroscopy, E. Sutherland and coauthors showed that the ratio of ribose, adenosine and phosphate in this compound was 1:1:1. Later on, by inserting synthetic cAMP (cyclic adenosine monophosphate) precursor, it was confirmed, that this was the second messenger for this reaction (Sutherland *et al.* 1968). Few years later, second cyclic messenger cGMP (cyclic guanosine monophosphate) was discovered.

At that moment most of the research was centered on Eukaryotic kingdom of life. However, (Ross *et al.* 1987) showed, that in Prokaryotic organisms similar processes can take place. For the formation of cellulose by *Acetobacter xylinum*, allosteric cofactor c-di-GMP (cyclic diguanosine monophosphate) was produced by the bacteria. Unfortunately, for decades it was completely ignored by microbiologists. Later it was shown, that this compound plays key role as second messenger in a lot of various (biofilm formation, motility, virulence, the cell cycle, differentiation, etc.) processes in different bacterial species (Romling *et al.* 2013; Krasteva *et al.* 2012).

In 1993, cAMP was also shown to function in bacteria as well as ppGpp (tetraphospho-guanosine) (Potrykus and Cashel 2008; Kolb *et al.* 1993) Until 2008, c-di-GMP was the only cyclic dinucleotide known. The novel c-di-AMP (cyclic diadenosine monophosphate) was found together with its cyclase protein (DUF147, later named as DAC), nevertheless the function remained unknown. Later it was suggested to play role in DNA damage control during cell-cycle checkpoints (Romling 2008; Bejerano-Sagie *et al.* 2006; Witte *et al.* 2008; Romling *et al.* 2013).

Last key point emerged quite recently. Novel asymmetric molecules, c-AMP-GMP (Cyclic guanosine [3`→5`] / [2`→5`] monophosphate–adenosine [3`→5`] monophosphate), were found in mammalian cells and as a product of *Vibrio cholerae* metabolism. The function of the first one is to control bacterial and viral infection via binding (as well as other known c-di-NMPs) to STING protein, which triggers interferon response, while the second takes part in pathogenic pathways (Wu *et al.* 2013; Gao *et al.* 2015). The structural inequality was studied by NMR and all three currently known proteins for c-AMP-GMP synthesis produce only 2`3` structure and only very recent studies reveal first 3`3` cyclase (Gao *et al.* 2013; Hallberg *et al.* 2016).

Thus their recent identification, the c-di-NMPs could be a huge piece of the puzzle, which is the metabolism of Prokaryotes.

1.2. Cyclic dinucleotides structure and mechanism of functioning

The discovery of each of the currently known c-di-NMP was serendipitous and, at least in case of c-di-NMP, quite underestimated.

Cellulose biosynthesis in *A. xylinum* (currently *Gluconacetobacter xylinum*) was thought as a simplified model of this process in plants. Surprisingly, the purified cellulose synthase showed much lower activity, in comparison with whole untouched cells as well, as whole membrane fraction. That suggested that there was some cofactor inside the cell or in the membrane. Purified protein crystal contained some specific form of guanosine-phosphate at ligand-binding site. Later, using NMR, Ross *et al.* (1987) showed, that both synthetic c-di[3`5`]-GMP and natural ligand had the same coupling constant and all the peaks were the same. Its structure is shown on Figure 1.

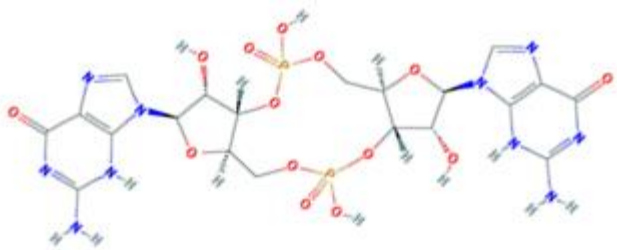


Figure 1. Structure of c-di-GMP. The structure resembles two ribo-guanosine monophosphates, linked with each other by phosphate groups in position [3 →5] of the cyclic sugars, the molecule is symmetric, it has two phosphate groups and two amino groups, which can be easily ionized (National Center for Biotechnology Information 2016).

Unlike linear oligonucleotides, the compound is stable in pH range of 3-10 (standard conditions, tested for 1 hour) and it was thermally stable (100°C), and did not degrade in human serum after 24h (Hyodo and Hayakawa 2008). That could be explained by mutual protection of the phospho-group via its linkage to 5` position on the paring ribose. This lead to low flexibility and, consequently, prevents from attack on the 2` OH group.

Another feature of c-di-GMP is its spontaneous stacking. By now it is found that the compound can form dimers (Fig. 2.), stabilized by H-bonds, which also takes part in signal transduction; tetra-, octa- and higher oligomers are reported, however their biological function remains unknown (Zhang *et al.* 2006).

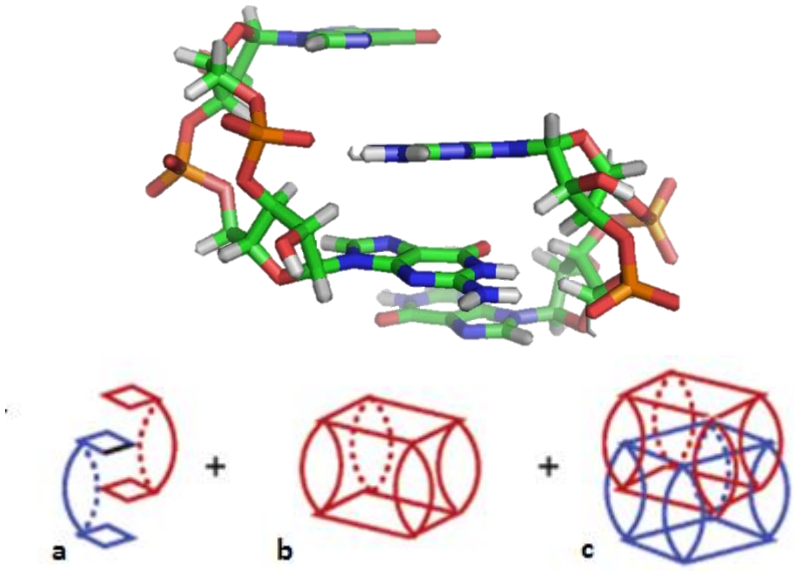


Figure 2. Different non-covalent staking forms of c-di-GMP. C-di-GMP dimer form. Carbon atoms shown in green, Nitrogen – blue, Oxygen – red, Phosphorus – orange and Hydrogen – grey (Romling et al. 2013). Oligomeric forms of c-di-GMP: a – dimer, b – tetramer, c – octamer. Both forms are shown to be present in natural and synthetic samples (Zhang et al. 2006).

Unlike c-di-GMP, other natural compounds of this class have not been reported to form any oligomeric structures. For c-di-AMP, it is considered to be due to the lack of oxygen in the second position of nucleobase and for c-AMP-GMP it is speculated to be linked with the

asymmetry of both nucleobases and phosphate linkage. Such asymmetry between phosphate and ribose in c-AMP-GMP molecule was quite controversial: original discovery paper reported it to be symmetrical [3` →5`] molecule, however all the NMR spectra of the purified molecule showed additional peaks, which could be explained either as a mixture of two symmetrical isomers or on asymmetrical. The latter was confirmed by additional investigation (Gao *et al.* 2015).

Specific structure of the molecule requires very distinct enzymes for both natural synthesis and degradation. These enzymes have specific motifs in their active sites, which react only to the molecule or its precursor with high specificity, though, some side products are demonstrated (Schirmer and Jenal 2009; Paul *et al.* 2010). It is also quite frequent, that the protein has both domains – GGDEF (cyclase activity) and EAL (phosphodiesterase activity) – on the single molecule. However, the combination of GGDEF and HD-GYP – which is another phosphodiesterase domain – is not reported (Fig.3.). Such construction suggests, that effects of the compound should be localized, which goes well with the diversity of the processes, c-di-NMPs are involved in (He *et al.* 2016; Krasteva *et al.* 2012).

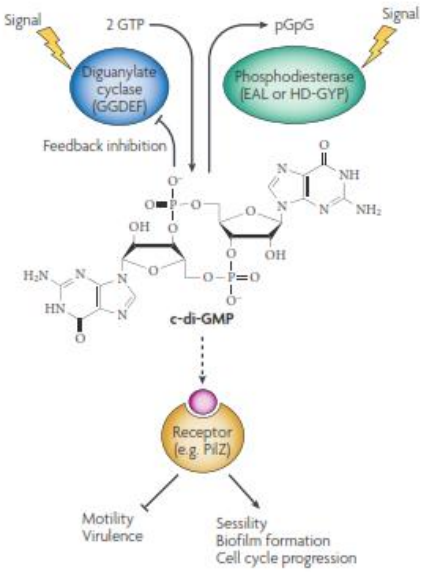


Figure 3. Generalized scheme of c-di-NMP pathway (example of c-di-GMP). The precursor is always the molecule of one class – 2 NTP, which through 2-step reaction is turned to corresponding cyclic dinucleotide. In of PDE (phosphodiesterase), there are two possibilities: EAL-domain will break one [3`→5`]-bond, while HD-GYP will go further and produce two separated GMPs (Schirmer and Jenal 2009).

These domains are highly conservative between different bacterial species and STING proteins in mammals, suggesting that it has significant importance for the bacterial metabolism and adaptation abilities (Gao *et al.* 2015).

As to main targets of c-di-NMP, besides EAL/HD-GYP degraded domain proteins they are PilZ-domain proteins and riboswitches – small specific RNA molecules. PilZ domain proteins were predicted directly upon discovery of c-di-GMP (Romling *et al.* 2013). They are transcriptional factors and were described in various bacteria, including *E. coli* (flagellum functioning, biofilm formation), *C. crescentus* (flagellum functioning), *V. cholerae* (biofilm formation, virulence) (Römling 2002; Tischler and Camilli 2005). As to riboswitches – a specific family of this RNAs, that can modulate gene expression, when bound with ligand – was found to react exclusively on c-di-GMP presence, even *in vitro*, which is a rear case. The sequences are reported to be upstream of *pde* or *dgc* genes' sequences that suggests their involvement in feedback control of the CDN amount. Unfortunately, mechanisms, in which these riboswitches are involved are still unknown, but presumed, that they could be involved in stress response (Romling *et al.* 2013; Sudarsan *et al.* 2008; Kalia *et al.* 2013).

On the contrary, mammalian c-AMP-GMP has specific target – STING protein that reacts to all the known c-di-NMPs, but in different site. The pathway is interesting in the point, that the compound is distinguished from other CDNs of bacterial origin. Mammalian c-AMP-GMP if produced as the response to free double strain DNA in cytoplasm of cells (Sun *et al.* 2013; Hallberg *et al.* 2016).

For development and further identification, it is of great importance to have chemical methods and instrumentation. Biological methods were and still are useful to get functional information of the ligand and targets, as well as pathway discovery. However, when it comes to structure identification and confirming the identity of compound, the chemical methodology is incomparably more advantageous.

1.3. Reported methods of identification of cyclic dinucleotides.

Many instrumental and non-instrumental methods have arisen in recent years and most of them are targeted for c-di-GMP. LC-MS is still the most widespread method for identification of the substance. However, some properties of c-di-GMP allow to identify its small amount in solution with help of titrimetry. The method, based on the thiazole orange property of fluorescence, which is remarkably enhanced in complex with quadruplex of c-di-GMP, was developed previously (Nakayama *et al.* 2011). Another method, also developed by this group, was the use of hemin and proflavine to aggregate c-di-GMP with Peroxidase. This gives a way to use simple colorimetry for the molecule. Such methods showed quite low detection limit of 5 and 1 μM respectively. These methods are quite robust and fast, and, what is of high importance, they are particularly specific for c-di-GMP. There was no report,

concerning detection of other CDNs, that combines with the fact of no reported non-covalent interactions of other molecules of the class (Nakayama *et al.* 2012).

Another complexing-based semi-quantitative titrimetric analysis method was developed with the help of specific diamidinium/iminium aromatic compounds in presence of K^+/Li^+ as complexing molecules. The results were measured using CD-spectra of initial solution and solution of complexes. Authors showed, that two molecules of the tested group (diminazene actuate and auramine O) incorporate at least 50% of c-di-GMP from solution and the presence of c-di-GMP can easily be detected by the method (Kelsey *et al.* 2012). The drawback is CD of complex samples and, even if to use another method for detection, the complexing may lead to further polymerization of c-di-GMP, forming chains of tetramers.

One more sophisticated way, is incorporating above mentioned riboswitches, that was shown by Zhou *et al.* (2016). For the investigation, they used natural I type riboswitches for control of *lac* operon in *E. coli*. Such riboswitches use two molecules in (trans-) dimer configuration, shown above. Authors showed, that use of one modified c-di-GMP-dependent riboswitch gave 16% difference between control and mutant strains in galactose catabolizes after addition of c-di-GMP. Incorporation of two riboswitches to regulate *lacZ* gave up to 50% difference. While being highly specific and having low detection limits, method is not enough reproducible and is too complicated for routine analysis.

As to LC-MS methods, the most widespread is one, described by Spangler *et al.* (2010), which is set as well established and specific enough for use in biological research. The method was used for c-di-GMP identification and quantitation in *E. coli* and some other bacteria. For LC they used C18 pyramid column with several column guards. Eluent A was water with 10mM ammonium acetate and 0.1% acetic acid and eluent B – pure methanol. For elution, they used 100% A to 70% A linear gradient. For MS/MS API 3000 with ESI was used in positive ion mode. The following transitions were monitored: +691/152, +691/135 and +691/248 (for confirmation). As a result, method is capable of detecting c-di-GMP in low ng/ml concentration and perfectly works at 2 – 8 μg/ml range. All the fragments where important in order to distinguish from interferences. For internal standard cXMP (cyclic xanthine monophosphate) was used.

Alternatively, Walker and Berkowitz (2013) proposed a method, which was able to separate most of known interfering compounds (including open cycle analogs) already by chromatography. This method was found to be suitable for both c-di-GMP and c-di-AMP. During bacteria growing, they performed the extraction steps two times (second time without hitting to 95°C (for proteins denaturation) with cooling in between for 20 minutes to 4°C,

while previous study suggested 1 time. By doing this, they got rid of most of the protein fragments and made the sample suitable for storing without losses at -20°C. Here they used acetonitrile for the eluent B. They also showed decrease in retention time without separation loss in HPLC. For IS $^{15}N^{13}C$-c-di-GMP was used.

In addition, method for LC/Q-TOF was also developed. It was used as a control for the riboswitch-based assay, described previously by Zhou *et al.* (2016). For HPLC they used Agilent™ C18 reversed-phase column (specific name not given) with isocratic elution (98% of water (with 0.2% of ammonium acetate and 0.1% of acetic acid) and 2% of methanol). For MS measurement, they used scanning mode with range *m/z* 100 to 1000 with scan rate of 2 spectra per second. Even with all the simplicity of method, authors were able to get accurate results.

As far, as research on c-GMP-AMP is only in the early phase, most of the authors use qualitative approach and much less precise instruments, like in the research on the structure of c-GAMP (Gao *et al.* 2015). Scientists used simple system of HPLC coupled with single quadrupole mass spectrometer. The eluents were water (with 10mM ammonium acetate and 0.1% acetic acid) and acetonitrile. They report the elution of the compound between 20 and 30% of acetonitrile (column – ODS-3 C18). In MS specific predicted mass was used (-673). They showed, that the 3`3` standard substance elutes later and only compound with same mass and retention time eluted from biological samples and protein assays.

Concerning modifications, which were used for research on each of the CDNs, a question of creation of a ubiquitous method for all combinations of nucleobases arises. For solving this question, one should first find these compounds in nature and synthesize them as pure as possible.

2. MATERIALS AND METHODS

2.1. Synthetic c-di-NMP and their properties

Main material, used for the investigation, c-di-NMPs (or CDNs), are compounds, which have a structure of two nucleotides, symmetrically or asymmetrically linked by their phosphate groups to ribose. In their turn, nucleotides incorporate a molecule of sugar (in nature ribose or 2-deoxy-D-ribose), phosphate groups, linked to its fifth carbon and a so-called nucleobase, which is linked to its first carbon. The nucleobases are specific heterocyclic compounds that belong to either purines or pyrimidines. Natural known bases are shown in Figure 4. Adenine, guanine and cytosine are main components of RNA and DNA, while uracil is specific for RNA and thymine is DNA component. Other bases are really found in specific RNAs, except for inosine and xanthine, which are precursors of adenine and guanine. Normally, nucleotides are connected via singular bond between ribose hydroxyl group, connected to third carbon and phosphate at fifth carbon position. In such manner the linear polymers are formed. However, in case of dinucleotides, both phosphate groups of two-nucleotide molecules attack hydroxyl of each other and such way form cyclic structure. That is the process, that happens during two steps protein-mediated reaction (Hyodo and Hayakawa 2008; Walker and Berkowitz 2013).

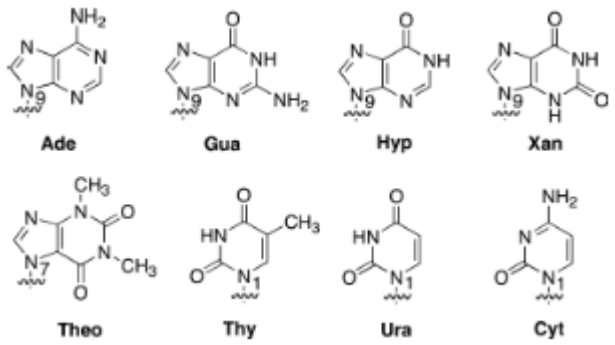


Figure 4. Natural nucleobase residues: adenine, guanine, hypoxanthine, xanthine, theophylline, thymine, uracil, cytosine. Most of these compounds in nature have different modifications for specific purposes (Clivio et al. 2013)

In organic synthesis, there are different strategies of forming such compounds. The most used are the following three (depending on initial compound): free nucleotide (Hyodo and Hayakawa 2008) or nucleoside-containing compound, for instance isobutyryl-TBS-protected nucleoside phosphate (Gaffney *et al.* 2010) or nucleoside phosphoramidite (Hayakawa *et al.* 2003). Compounds synthesis for our experiment was based on second variant, as the easiest; however, the initial compound was not protected. It incorporated eight synthesis steps and two purification steps. Another difference from mentioned method is that

HPLC was used for last step, instead of crystallization to avoid possible degradation of obtained compounds, as their properties could differ from those of known compounds. For each of the compounds NMR spectra were acquired to check the purity and confirm the structure of synthesis products. All of them showed no significant impurities and had structural peaks, as predicted (data not shown).

All the ten compound properties were examined. The compounds were found to be stable under broad range of pH (2.3-8.5) and relatively stable in water (40% degradation of c-di-CMP after 12 hours at room temperature). Novel compounds did not show the ability to form oligomeric structures, like was shown for c-di-GMP, except possibly c-di-CMP.

2.2. Laboratory equipment and reagents

As all stages of experiment were carried out in one lab, we had all the tools and reagents necessary for them.

Instruments:

- Chemistry-HYBRID pump RC 6™ coupled with Bachofer vacuum concentrator centrifuge (Vacuumbrand®, Wertheim, Germany)
- Ecotron™ shaker (Infors AG, Bottmingen, Switzerland)
- Centrifuge 5427 R (Eppendorf®, Hamburg, Germany)
- Centrifuge 5810 R (Eppendorf®, Hamburg, Germany)
- Vortex™ mixer (VWR International bvba, Leuven, Belgium)
- Heraeus Biofuge Pico™ (Kendro®, Hanau, Germany)
- Branson® Sonifier SFX250™ (Emerson Electric Co., St. Louis, US)
- Liquid chromatography unit including SPD-20A™, DGU-20A™, 2 LC-20AT™, CBM-20A™ (SHIMADZU®, Kyoto, Japan)
- Mass spectrometer LCMS-2020 (SHIMADZU®, Kyoto, Japan) with ESI interface
- Column heater THERMASPHERE™ TS-130 (Phenomenex®, Torrance, US)

Tools:

- MICROLITER™ syringes 10 and 25 µl (Hamilton Bonaduz AG, Bonaduz, Switzerland)
- 0.2, 20, 200 and 1000 µl Pipettes (Eppendorf®, Hamburg, Germany)
- 1.5, 2.0 and 0.5 tubes (Eppendorf®, Hamburg, Germany)

Reagents and buffer components:

- Acetonitrile HiPerSolv CHROMANORM™ (VWR CHEMICALS®, Fontenay-sous-Bois, France)

- Ammonium formate HiPerSolv CHROMANORM™ (VWR CHEMICALS®, Fontenay-sous-Bois, France)
- Formic acid Optima™ LC/MC (Fisher Chemical®, Illkirch-Graffenstaden, France)
- Water, purified with Milli-Q Direct 8 system with 0.22 µm Millipak Express 40 Final Filter (Millipore SAS, France)

Other compounds were of analytical purity grade.

2.3. Culture growing

The main focus was put on *Escherichia coli*, ssp K-12 variants (MG1655, BW). These strains are interesting in the point, that MG1655 preserves very high amount of genetic material richness and this makes it the closest to *E. coli* WT (wild type). Second variant, BW25113 has one of the richest gene knockdown collection, including part of the genes of interest (of DGC-candidates and CDN PDE-candidates families) (Baba *et al.* 2006). This bacterium was chosen not only because of its well-studied proteome and genome as a model organism, but also because it is shown to have ability for c-di-GMP production, while induced for biofilm formation (Caly and Bellini 2015; Krasteva *et al.* 2012).

Besides model organisms (*E. coli* for gram-negative and *B. subtilis* for gram-positive), we used other bacteria species, which were reported to produce either c-di-AMP or c-di-GMP. To check, whether test medium can bring artifacts, we used eukaryotic yeast species (*S. cerevisiae*). All the data about organisms, used for the experiment, are in table 1. We were not able to check the most known producer of c-GMP-AMP, *Vibrio cholerae* as the lab does not have permission to work with hazardous species.

Table 1. List of organisms, reported to have CDNs

Bacterium name	Strain	Growing media	Known CDNs	Reference
Escherichia coli	K-12 MG1655 K-12 BW25113	LB, M9	c-di-GMP	(Romling *et al.* 2013)
Bacillus subtilis	DSM10	LB, MSgg	c-di-AMP	(Romling 2008)
Saccharopolyspora erythraea	ATCC 11635	Medium 65	c-di-GMP	(Krasteva *et al.* 2012)
Staphylococcus cohnii	GH137	LB	c-di-GMP	(Romling *et al.* 2013)

Bacterium name	Strain	Growing media	Known CDNs	Reference
Comamonas testosteroni	KF-1	LB	c-di-GMP	(Francis and Corbin 1999)
Pseudomonas putida	KT-2440	LB	c-di-GMP	(Zogaj *et al.* 2001)
Caulobacter crescentus	CB15N	PYE	c-di-GMP	(Spangler *et al.* 2010)
Xanthomonas campestris	ATCC 33913	Medium 54, TY	c-di-GMP, possibly c-di-AMP	(Hyodo and Hayakawa 2008; Huynh and Woodward 2016)
Saccharomyces cerevisiae	BY4741	YPAD	-	

For composition of media – see annex A

Generally, the procedure of growing was as follows: bacteria were inoculated in corresponding nutritious medium and left to grow for time, approximately meeting the needed living cycle phase (during the process each hour the optical density at 600 nm was measured until it reached correct level for the phase).

2.4. Sample preparation

After reaching desired OD, sample was separated in several (usually 5) portions and processed, using the following protocol, based on (Spangler *et al.* 2010) with significant modifications, therefore all the steps are shown:

I. Centrifugation (4°C, 30 min, 1700 × g[1])
II. Resuspension and wash of each pellet with 5 ml phosphate buffer (pH = 7.4) and combining in one tube
III. Centrifugation (4°C, 30 min, 1700 × g)
IV. Resuspension of pellet in 6 ml H_2O (Milli-Q)
V. Lysis of cells by sonication (50% of maximum power, 6 min.)
VI. Sample for Bradford assay (200 µl)
VII. Transfer to 15 ml tube and addition of approximately equal volume of cold (4°C) acetonitrile
VIII. Centrifugation (4°C, 30 min, 1700 × g)
IX. Discarding pellet and freezing supernatant (in 2 ml Eppendorf tubes)

[1] G corresponds to standard acceleration due to gravity and equals to 9.81 m/s^2

X. Volume reduction by freeze-drying (approx. 1.2 h)
XI. Centrifugation (4°C, 30 min, 21380 × g)
XII. Transfer to "cut-off" filters
XIII. Centrifugation (4°C, 30 min, 1700 × g)
XIV. Complete freeze-drying (8-12 hours)

These steps could be separated in groups: bacteria processing (step I. to V.), sample nucleotides separation (VII. to X.) and final purification (XI. to XIII.). Lysis-connected steps (III. to V.) were usually repeated 2 times. Step VI. was done for more precise determination of amount of the bacteria, as OD_{600} can give only very approximate information due to various possible effects, including contamination. Next steps were performed to separate liquid fraction of small molecules from cell debris and big proteins, remaining in liquid phase. In final part is another purification, performed with specific filter, which is capable to clarify sample from most macromolecules, which could remain in the sample. For this, so-called "cut-off" filter is used. Such filter consists of two main parts: the tube (usually 15 or 50 ml with specific cap) and the filter, which is placed inside the tube, in its upper part. The filter is a membrane with pores of specific size, which is measured correspondently to smallest mass (in Daltons (Da) – 1 Da ≈ (1.66×10^{-27}) kg) of protein that can pass through it. For this step we used filter with the smallest pore size available on market (3 kDa, SIGMA-ALDRICH®), that roughly corresponds to 200 amino acids protein. The sample is transferred inside this filter and then centrifuged. We performed this step 2 times. Upon completing this procedure, sample was completely dried by lyophilization and then one of the Eppendorf tubes, containing sample was use for LC-preseparation

2.5. LC/MS parameters and solutions

The central part of experiment was the separation and detection of 10 synthetic CDNs with the help of liquid chromatography-mass spectrometry.

To have better and more stable method, we used system of two columns ("2D-HPLC"), first of which was used to collect the entire dinucleotides fraction, separately from other components. To achieve this, we choose Clarity™ Oligo-RP column (PHENOMENEX®) with precolumn. The column is designed to distinguish between oligonucleotides with even single base change, as well as between single and double stranded (ssDNA/RNA and dsDNA/RNA) in high concentrations of acetonitrile. We used relatively high aquatic conditions (88% at the elution) to reduce such discrimination, but enhance differentiation between compounds with different amount of nucleobases (especially cyclic

mononucleotides, which can affect separation on following columns). In addition, such purification allowed to separate CDNs from most of other compounds and, therefore, prolonged functioning and enhance separation (Fig. 5).

For the separation of nucleobases, two columns were used. Main routine measurements were carried out, using C18 GravitySB™ (NUCLEODUR®) with particle size of 3 µm (MACHEREY-NAGEL GmbH & Co. KG 2015). The column is based on standard C18 with octadecyl modification and extensive endcaping. Such design should let the column to be able to distinguish among early-eluting polar compounds with close properties, including nucleotides, however it is not good in distinguishing differences among rotational or other steric isomers (Fig. 5).

Second column was used for confirmation of new compound presence in sample. This column (C18 Isis™, NUCLEODUR®), also with 3 µm particles, is highly enriched with cross-linking and specific polymeric modifications (MACHEREY-NAGEL GmbH & Co. KG. 2015). Such modification gives to column good distinguishing properties towards the different steric isomers with same atomic composition. Using both of the columns, we suppose to see with high probability only the peaks, which correspond either CDNs or their products and not the linear dimers (product of RNA and CDNs partial degradation).

A **B**

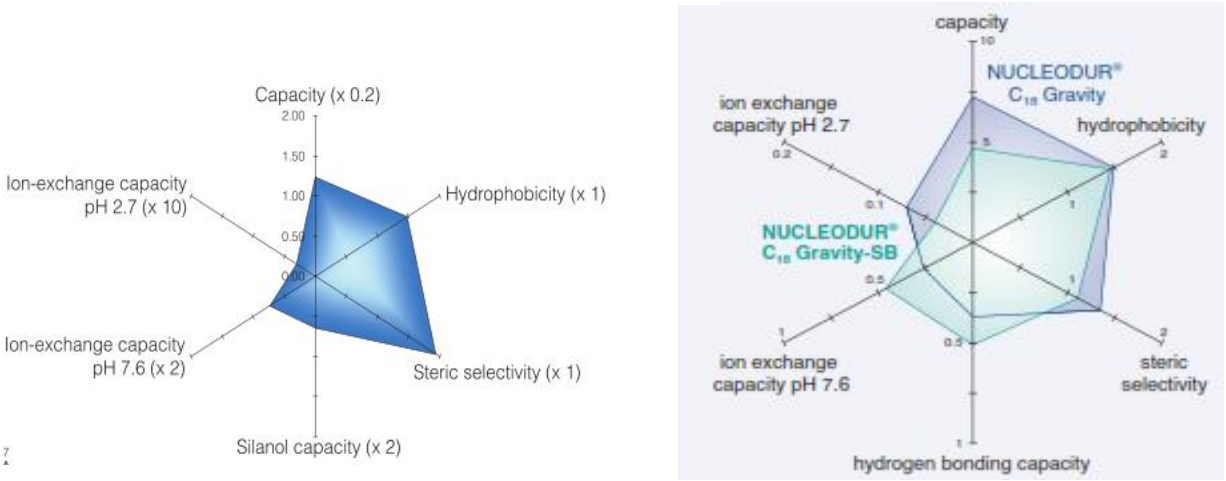


Figure 5. Tanaka plots of C18 Isis (A) and C18 Gravity-SB (B). The difference between properties of the columns give much more narrow range of possible compounds, which can elute simultaneously with target species (MACHEREY-NAGEL GmbH & Co. KG 2015; MACHEREY-NAGEL GmbH & Co. KG. 2015)

To perform the LC steps, all the time the same eluents were used (for purification and for LC-MS measurement). We used system of two buffers of the following content:

- Buffer A – H_2O (Milli-Q™), with addition 0.1% formic acid and 10 mM ammonium formate;

- Buffer B – acetonitrile with addition of 0.1% formic acid.

All columns were stored with same storing solution, consisting of 80% of acetonitrile and 20% of water (Milli-Q). Before the measurements, each column was flushed with working buffers and time program was performed twice – firstly, without adding any sample and secondly with adding of 10 synthetic CDNs. Once a week both analytical columns were checked to have the same pressure at parameters as shown in their certificates. After using, all of the columns were washed with the initial solution for 10 column volumes.

Most of the liquid chromatography parameters differ during the use of above-mentioned columns. For Clarity column, the injection volume was 20 µl (that represents 20 ml of initial culture) and the flow rate was constantly kept on 3.5 ml per minute. Pressure was controlled not to exceed 140 bars. No temperature control was used. Program started with 0% of B till 1.5 minutes, followed by gradient till 1.6 minute to 12% of B with isocratic elution till 4.5 minutes. Re-equilibration of the column was achieved by washing with 95% of buffer B from 5 till 7 minutes. The detection was carried out by UV-detector at wavelength 260 nm and sampling rate – 2 Hz. Fraction of dinucleotides was collected and completely freeze-dried.

After finishing lyophilization, 20 µl of buffer A were added to the sample. For analysis with C18 GravitySB and C18 Isis columns, 1 to 4 µl (representing 1 to 4 ml of bacteria culture) of sample were injected per run, depending on whether the sample was previously purified with Clarity column and on total current spectrum after first injection of 2 µl -. For C18 GravitySB, the flow rate was 0.3 ml/min, temperature was fixed to be 30°C, and the pressure was visually controlled not to be higher than 200 bars. As in previous method, time program started from 0% of B with rapid jump to 4% at 0.1 min, followed by two step slowing of gradient: to 6% at 2 minutes and to 8% at 5 minutes. The wash with 95% B from 7 till 12 minutes was performed to refine the column properties and wash out all the contaminating compounds of the sample.

For C18 Isis column flow rate was 0.2 ml/min and temperature of 30°C, pressure not exceeding 180 bars. Program started at 0% of buffer B, quickly raised to 4% at 0.1 min and had 3 linear gradient parts: 4% to 6% at 2 minutes, 6% till 7% at 4 minutes and 7% till 20% at 9 minutes with column wash till 18 minutes at 95% of B. In both of latter and the previous programs, when measuring biological sample combination of 10 synthetic CDNs in concentration of 5mM of each was used for spiking the samples. These compounds were added in either 1 or 2 µl of solution. After performing the program 2 to 4 minutes were let for the column to return to initial pressure values (174 bars for C18 GravitySB and 169 bars for C18 Isis)

Mass spectrometer was used only with the two latter columns and was coupled with HPLC system via ESI interface. Drying and colliding gases were N_2 at 8 and 1.5 L/min respectively. Interface and heat block temperatures were set to 250°C and 200°C respectively. Detector voltage was 1300 V, vacuum – $4.8 \cdot 10^{-4}$ Pa. Preliminary measurements were performed in scan mode, while actual analysis – in SIM mode for higher signal/noise ratio. Both Negative and positive ionizations were performed

For samples, separated with C18 GravitySB, actual recording of mass spectra was performed from 3 to 6 minutes and by C18 Isis – from 5.5 till 10.5 minutes.

2.6. Data treatment and processing

All of the measurements and data processing were performed using standard tool – Shimadzu LabSolutions™, which contains applications for HPLC and LC-MS experiment performing, as well as tools for viewing spectra and peaks' manipulation ("Browser" and "Postrun"). Most of the spectra area and height data were acquired in automatic mode with following parameters (separately for each *m/z* value): maximum of peaks – 20, slope – 1200/s (for c-di-CMP – 700/s), broadness – 0.07 min. In cases of obvious peak overlap or bigger peak observed in the sample (particularly for c-di-UMP), manual modifications were made. All the calculations and graphical material preparation were performed using Microsoft© Excel™ 2016.

Calibration curve was built for the reasonable range of 1.25 to 25 mM. Stability and reproducibility estimated on basis of measurements of standard mixture over 4 months for C18 GravitySB and about two months for C18 Isis.

Reproducibility evaluations contained 11 measurements for Isis and 31 for GravitySB, performed at different days at the same conditions and from the same set of standard compounds. Measurement with C18 GravitySB started from late February and for C18 Isis – beginning of April 2016.

Each of the biological samples was measured in two duplicates – twice only sample and twice a mixture of same amount of sample with standard. For repeatability, samples from two independent cultures were measured. Values were calculated on basis of mean of all four same measurements, while standard deviation – on basis of averages of two independent measurements. Calculation of approximate concentration were performed on basis of relation of sample value-to-value of sample plus standard. Finally, results were given in nM and ng/ml of CDNs in initial volume, normalized to OD_{600} of culture just before sonication.

3. RESULTS AND DISCUSSION

3.1. Method performance

In this section, data of the method optimization and control are presented. As the use of synthetic standards was throughout the whole experiment, all the methods were optimized for best separation of these compounds.

3.1.1. Purification

Separation of the dinucleotide fraction from other compounds, active at 260 nm wavelength, was performed at the 12% of acetonitrile phase of the chromatographic program. For testing, mixture of CDNs was spiked with 20 μM of mononucleotide Adenine (Sigma-Aldrich, Hungary) and nucleotide oligomers (Glen Research, US). This is shown on the Figure 6.

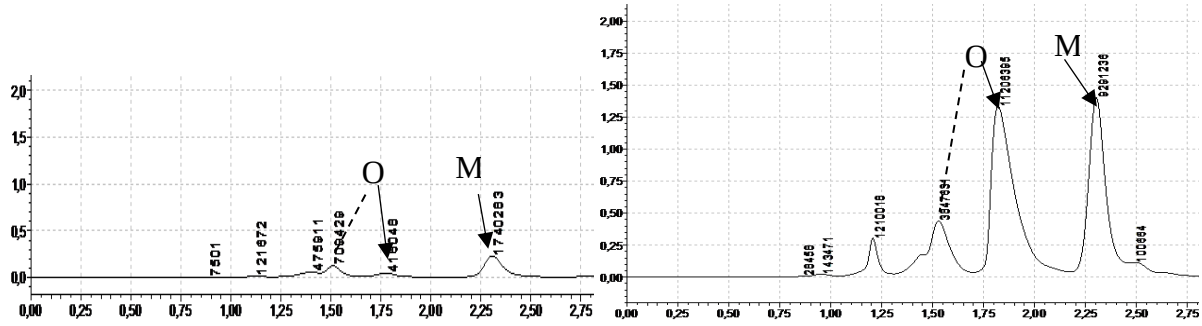


Figure 6. Spiking of standard for Clarity Oligo-RP purification test. First chromatogram corresponds to non-spiked sample, while second is spiked. O – oligonucleotides, M – adenosine monophosphate. Dashed line directs to unknown peak, which may be impurity in commercial oligomer or result of formation of short double strands of RNA. Numbers correspond to peak areas Part after 2.8 minutes is removed

Method was optimized to be highly stable in the elution times even without temperature control, with the fraction collected from 2.75 till 3.35 minutes (deviation less than 0.025 minute), which corresponds to approximately 2 ml eluent solution with compounds. Although, peak did not have symmetric shape and had tailing (Fig. 7.), it was sufficient for such purification. For standard, the peak height was around 4000-5000 intensity units.

Unfortunately, this method does not allow collecting full fraction of c-di-CMP, which was eluted earlier (2.6-2.65 minutes) and only around 30% remain in the main fraction. Such behavior suggests impurities and possible degradation of the compound.

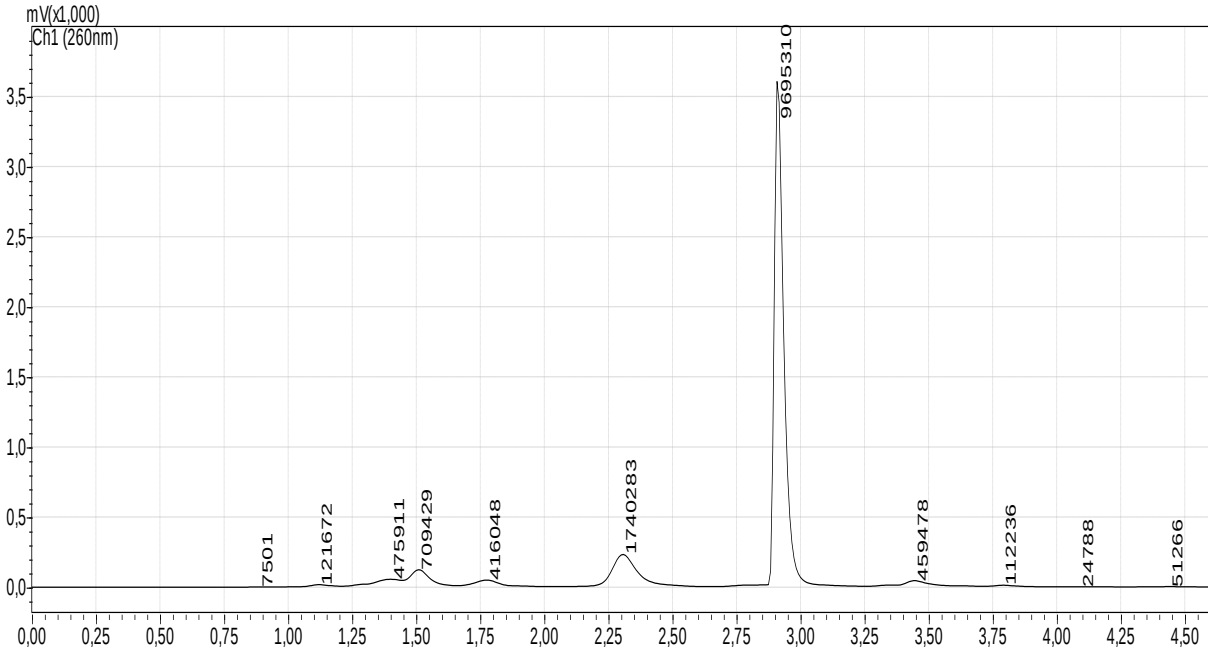


Figure 7. Chromatogram of standard during the purification. Most of the 260 nm- active compounds were eluting in the 12% acetonitrile, except for a minor peak, eluting just after the collected fraction. Numbers correspond to peak area. Part from 4.5 till 7 minutes is removed for better visualization; no peaks occurred in it.

After such purification, in combination separation with 0.22 µm filter, bacterial extracts were pure enough to observe changes at nanogram per milliliter level, in the positions corresponding to standards' retention time.

3.1.2. Separation of cyclic dinucleotides with different C18 columns

Columns used for LC-MS analysis, showed differences both in the order of the compounds elution and in overall retention time. In the beginning, samples were also measured without purification procedure for additional testing. Later it was abandoned, due to high column damage as one of the C18 GravitySB columns was destroyed. To stabilize retention time, columns were kept in the column heater at stable temperature 30°C. C18 Isis showed to have better retention time stability (Table 2) and higher distinguishing ability for c-AMP-CMP, however peak shape was much worse, as with C18 GravitySB and did not change much with significant changes to the gradient.

Both columns had their beneficial properties: the C18 Isis has better resolution and peaks' retention time fluctuates much less and C18 GravitySB had better peaks shape and re-equilibrates faster, so more suitable for quantification and every-day use. Still, both columns give artifacts, which can be rejected only by measuring sample on both of them.

All the CMP-containing compounds show to have an additional peak, coupled closely with the main one on C18 Isis. This is possibly due to synthesis non-selectivity, causing

standard to be a mixture of several CDN optical isomers or spontaneous transition between isomers in the mixture, though the NMR tests did not show it for most of the compounds and only slightly for c-di-CMP.

We decided to calculate the approximate amount only using C18 Gravity SB column and confirm the possible peaks using C18 Isis.

Table 2. Cyclic dinucleotides properties

Name of compound	Mass, Da	Retention time C18 GSB, min	Peak shape C18 GSB	Retention time C18 Isis, min	Peak shape C18 Isis
c-di-CMP	≈610	3.79 ± 0.09	Broad, fronting	6.15 ± 0.1	Very broad, strong tailing
c-CUMP	≈611	3.83 ± 0.04	Good	7.75 ± 0.02	Good, minor tailing
c-di-UMP	≈612	4.16 ± 0.19	Good, two major peaks	9.64 ± 0.01	Broad
c-CAMP	≈634	3.92 ± 0.19	Good, minor fronting	7.66 ± 0.03	Strong tailing
c-UAMP	≈635	3.96 ± 0.12	Good	8.86 ± 0.02	Good, tailing
c-CGMP	≈650	3.94 ± 0.19	Good	7.81 ± 0.03	Strong tailing
c-UGMP	≈651	4.13 ± 0.19	Good	9.42 ± 0.02	Tailing
c-di-AMP	658.412	4.12 ± 0.17	Broad, minor fronting	8.67 ± 0.04	Tailing, minor broadening
c-AGMP	672.395	4.1 ± 0.22	Good	8.93 ± 0.04	Good, tailing
c-di-GMP	690.411	4.29 ± 0.22	Good	9.35 ± 0.05	Tailing

Approximate masses here are of the compounds, which were not found in nature. During analysis with mass spectrometer, their m/z ratio was set as x.15, where x – mass from the table.

To estimate lowest possible concentration and the linearity of method, calibration curves (Fig. 8) were built separately for each compound (annex C). The lowest linear amount was 2.5 nM (≈ 1.5 ng/ml) and method remained linear till 25 nM (≈15 ng/ml). These concentrations are lower, than in previously suggested method (Walker and Berkowitz 2013; Spangler et al. 2010). However, to calculate higher concentrations of CDNs, mentioned methods are recommended.

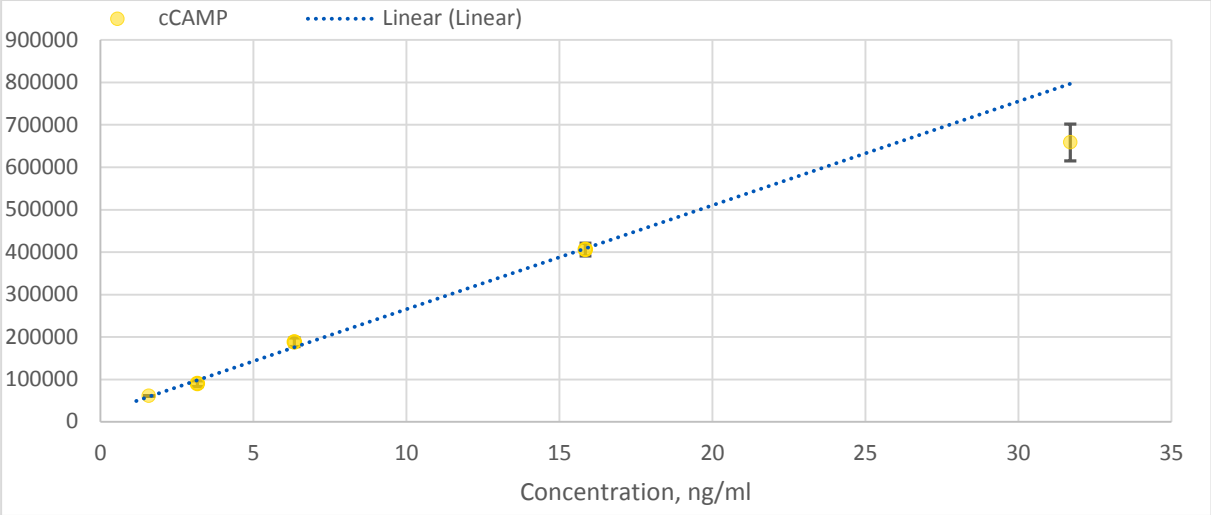
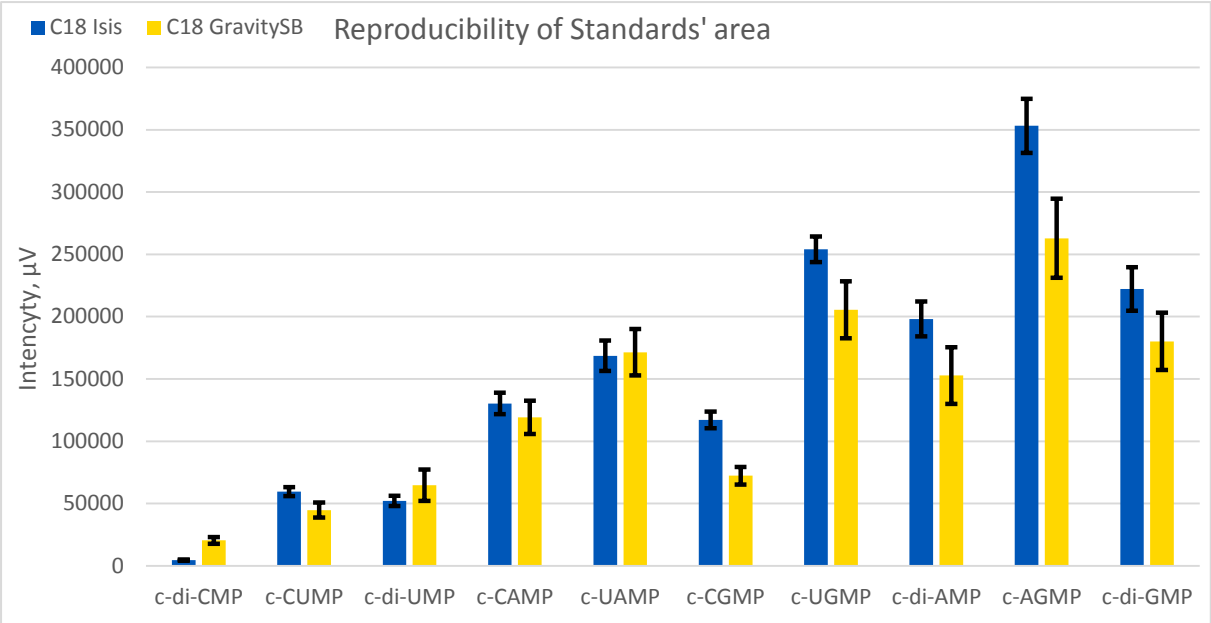


Figure 8. Calibration curve of the standard substances (on example of c-AMP-CMP) using C18 GravitySB. For lower concentrations, error bars are hidden being points.

Use of probes with lower concentrations appeared to have non-linear relationship of signal and the concentration was not estimated at that levels. As to qualitative amounts, the LOD in *E. coli* for samples each compound was between 0.12 and 0.02 ng/ml for C18 GravitySB and 0.05 to 0.01 ng/ml for C18 Isis. The only exception was c-di-CMP with respective values of 0.23 and 1.091 ng/ml that is due to the high losses of this compound during the LC purification. For detailed data – see annex B. As we put use both columns, overall limit for most compounds is of GravitySB and for c-di-CMP is the one of Isis.

Reproducibility (Fig. 9) on both columns is within 15% level (except of c-di-UMP on C18 GravitySB, which had 19.9%), and the fact, that first and last measurement difference is small (see annex D) that suggest that all samples, measured with the method remain comparable and the method has low short-time bias. Greater difference in c-di-UMP is due to the two peaks of the standard on GravitySB. This suggests that synthetic compound is the mixture of isomers with different hydrogen bonging properties, but not sterically different (one peak on Isis).

Measurements of synthetic standards were found not to be interfering with various known compounds of nucleotide nature, including various polyphosphate compounds, linear dinucleotides and cyclic mononucleotides. All of these compounds have different retention times either during Clarity Oligo-RP purification or during LC-MS separation and analysis with C18 GravitySB/Isis.



C18 GSB	13.5%	13.3%	19.5%	11.2%	10.9%	9.8%	11.1%	14.9%	12.0%	12.8%
C18 Isis	9.7%	6.1%	7.8%	6.6%	7.3%	5.7%	4.1%	7.1%	6.2%	7.9%

Figure 9. Reproducibility of standard compounds signals on different columns: blue corresponds to C18 GravitySB and yellow – C18 Isis. Error bards correspond to standard deviation of responses. Percentages below the Figure correspond to ratio between standard deviation and average area of above-mentioned molecule of standard mixture.

3.1.3. Results of mass spectrometric analysis

Ionization of compounds was performed in both positive and negative mode (under standard parameters for measuring). Signals in negative mode were more distinct and had better isotope profile (Table 3). This data was achieved from a direct comparison of average intensity of mass-ions in positive and negative modes.

In positive mode, nevertheless, was only one peak for c-di-UMP, instead of two in negative, when working with C18 GravitySB. On C18 Isis column negative mode gave only 1 major peak and 2 to 3 minor with intensity 4 or more times lower.

Table 3. Ionization in positive and negative modes.

Bases of CDN	**1**	C	C	U	C	U	C	U	A
	2	C	U	U	A	A	G	G	A
Ionization ratio	**pH 3**	2	20	20	5	3	5	7	6
	pH 7	10	25	100	5	20	5	50	6

Numbers here show, how much bigger the peak of the same compound is in negative mode, than positive.

Another problematic point with negative mode was appearance of doubly charged ions, which greatly lowered the detection ability for respective nucleotides (Fig. 10). This was most significant for c-di-CMP and at much lower level with other CMP-containing compounds.

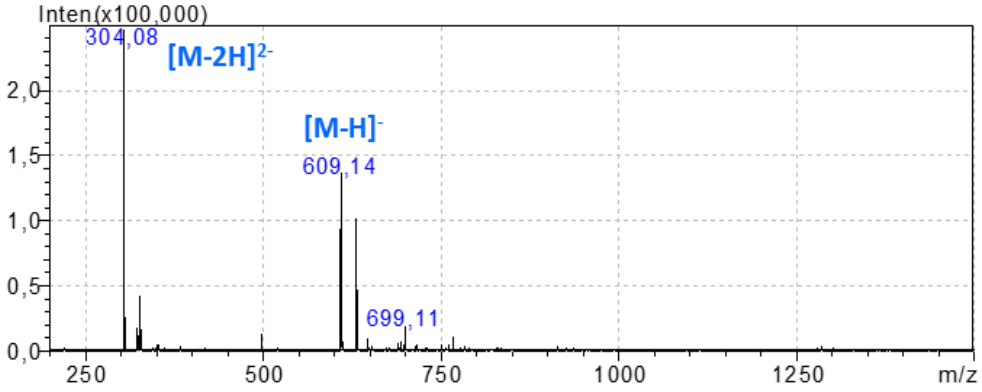


Figure 10. Fragment of mass spectrum in SCAN mode. Peak 609.14 corresponds to singly charged c-di-CMP and 304.08 – doubly charged

Throughout all of the analytical steps in the method, we were experiencing difficulties with c-di-CMP standard and possible candidates for it. This compound showed distinct properties from the rest and we were not able to incorporate fully it in the overall method. Easier double charging of c-di-CMP suggests, that this compound is either not purely synthesized, as was mentioned previously, or has distinct properties from other cyclic compounds, that does not clearly come from the structure. Latter may be, for instance, same constructs, as of c-di-GMP (Fig. 2) or another form of stacking, and is stronger, comparing to other compounds, and so effecting ionization. Another explanation would be that the compound is not stable and quickly degrades. But this is not deducible from the structural point of view.

3.2. Analysis of bacterial samples

The procedure, outlined in section 2.4. was applied to all the bacterial samples. The most crucial in terms of sample losses was bacteria destruction by sonication. Up to 50% of bacterial culture sample volume was lost during this procedure. However, passing standards though the whole sample preparation procedure did not show the significant difference in amount of compounds, suggesting, that neither of steps influence the compounds. This, nevertheless, may be different for compounds in cells, as they may be bound to some proteins, especially transmembrane, and be later stacked in cell debris. Therefore, this step needs significant optimization, when transferring to quantitative analysis.

Samples of bacterial extracts were measured without the bacteria name label (blind measurement). In the MS chromatograms were several big peaks in the same *m/z* ratio as compound, but they were not masking the regions of standards' and proposed candidates' elution.

Table 4. CDNs presence in various bacteria

Code	Organism	Phase	c-di-CMP	c-CUMP	c-di-UMP	c-ACMP	c-AUMP	c-CGMP	c-GUMP	c-di-AMP	c-AGMP	c-di-GMP
#15a	*St. cohnii*	exp			**yes**	**yes**						
#15b	*St. cohnii*	stat				**yes**						
#17a	*C. testosteroni*	exp		**yes**		**yes**				**yes**		**yes**
#17b	*C. testosteroni*	stat										
#18a	*Ps. putida*	exp							**yes**			
#18b	*Ps. putida*	stat										**yes**
#4b	*B. subtilis*	exp	**yes**			**2.77**						
#4c	*B. subtilis*	stat	**yes**									
#4d	*B. subtilis*	film	**4.52**		**2.8**	**yes**				**yes**		
#14a	*S. erythraea*	exp			**11.48**			**yes**				**yes**
#14b	*S. erythraea*	stat										
#16a	*S. cerevisiae*	exp				**yes**						**yes**
#16b	*S. cerevisiae*	stat										
#19a	*C. crescentus*	exp			**4.84**	**yes**						**yes**
#19b	*C. crescentus*	stat			**3.41**							
#20a	*X. campestris*	exp		**2.84**		**yes**	**yes**	**3.8**	**yes**	**yes**		**yes**
#20b	*X. campestris*	stat						**yes**	**2.9**		**yes**	
#20c	*X. campestris*	exp				**yes**			**3.45**	**yes**	**yes**	
#20d	*X. campestris*	stat							**3.17**	**yes**	**yes**	

All the found compounds' signals are referred as candidates. All the values represent the amount, measured with GravitySB (mean of two repeats of two independent cultures). If measured quantity was below the lowest of calibration points, but still clearly distinguishable, label "yes" was used to signify its presence.

Exp in this table stands for middle period of exponential growth phase and corresponds 3-5 hours of growing, Stat - stationary phase of bacterial growth and corresponds to overnight culture and film - biofilm growth, that was collected approximately after 10 hours of growing.

All the bacteria were grow on one medium, except of *X. campestries* and B. subtilis. First was in parallel on two media – samples 20a and 20b on medium 54 and 20c and 20d – TY-medium. Second was grown on specific medium to induce biofilm formation.

All the concentrations are given in nM

Table 4 shows the data on all the various bacterial extracts measured to test method. The compounds, which were found on one column, but not on both are not included (see annex E). We found, that we are able to detect all the previously identified compounds in different species at least in one stage of growth. The only exception was *Staphylococcus cohnii*, which we were no able to grow to high OD_{600} and latter attempts were not successful.

Beneficial decision was the usage of two columns for confirmation of the peak of the compound of interest. Measuring with a single column the purified sample did not give enough confidence, as in most samples there were non-reproducible peaks in independent cultures repeats.

Comparing, for instance data for *E. coli*, we were able to reject many peaks, which were not eluting at the same time, as standard (especially c-AMP-CMP candidate) on one of the columns.

Nature of the non-reproducible peaks, could be explained in many ways, but the most probable suggestions were artifacts from medium or differences in culture preparation (growing period, inappropriate medium preparation), since they were prepared not in the same time, but with several weeks difference.

To estimate losses and matrix effects in the LC-MS procedure, we took the ratio of average of difference between spiked and non-spiked samples to the average of standard (table 5) Results, show that various CDNs are affected differently, however, many of them have recovery around 85%. On the other hand, deviation of some sample peaks of compounds was up to 40% (c-di-AMP); this can be seen in table 5.

Table 5. Sample loss during the process of measurement on C18 GrasitySB

Name of compound	Direct standards' average	Samples, spiked with standards		Recovery
		Average	Deviation	
c-di-CMP	20444.73	15559.67	33.83%	76.11%
c-CUMP	44774.48	25921.67	2.84%	57.89%
c-di-UMP	64785.68	49413.00	28.81%	76.27%
c-CAMP	119151.71	78912.60	30.03%	66.23%
c-UAMP	171394.13	151307.00	23.40%	88.28%
c-CGMP	72385.50	64705.14	17.16%	89.39%
c-UGMP	205552.77	159472.75	20.43%	77.58%
c-di-AMP	152720.92	128118.93	39.31%	83.89%
c-GAMP	262913.65	215556.67	26.84%	81.99%
c-di-GMP	180203.77	155394.40	15.92%	86.23%

Finally, *E. coli*, was sampled at different phases of growing to test the hypothesis, that the production of CDNs may be dependent from bacteria growing stage. We prepared four samples, which should correspond to different points on the growing curve, labeled: early exponential (2-3 hours after inoculation), middle exponential (4-5 hours), late exponential (5-8 hours) and stationary (more than 8 hours after inoculation). Results could be seen in table 6.

Table 6. Growth-phase emerging of CDNs

***E. coli* variant**	**MG 1655**				**BW 25113**			
Phase	**exp early**	**exp middle**	**exp late**	**Stat**	**exp early**	**exp middle**	**exp late**	**stat**
c-di-CMP								
c-CUMP								
c-di-UMP								
c-ACMP								
c-AUMP								
c-CGMP								
c-GUMP								
c-di-AMP								
c-AGMP								
c-di-GMP								

Colors represent relative amount of corresponding dinucleotide. Pastel red – very low (< 0.2 ng/ml), pastel yellow – low (0.2 to 0.7), pastel green – middle (0.7 to 1.5) and deep green – high (>1.5). All the levels are based on approximate values, normalized to OD_{600} of the bacterial cultures before processing.

To control the influence of the medium on the response, salt-based medium M9 was used. This medium does not contain any of the biological extracts, like yeast extract in LB (see Annex A), therefore no medium effects on the signal are possible. In addition to c-di-GMP, we were able to confirm presence of c-AMP-CMP (1.08 ng/ml in middle exponential phase and 0.47 ng/ml in stationary), c-GMP-UMP (0.83 ng/ml in middle exponential phase) and c-GMP-CMP (0.29 ng/ml in stationary phase) candidates on it, as well as their dependence on stage of growing. This shows that the compounds come from the bacteria and not from the medium. However, c-CMP-UMP seems to come from the LB medium or is not produced in the used conditions.

Possible discovery of the cyclic dinucleotides, other than currently known, suggests that these compounds have even more diverse functions, than those confirmed. They may emerge as real analogs of complex secondary messengers' regulatory system in plants and animals, influencing all parts of bacterial life. Moreover, such diversity fit with the fact of varied

metabolism and environment adaptation abilities and limited compartmentalization of prokaryotic cell. As the compound candidates were found in most of the samples, they could be ubiquitous in the diversification signal strategy of all the *Prokaryota*.

Another important suggestion is the dependence of the amount of some of cyclic dinucleotides on certain phase of growing, while others were independent of it. If confirmed, this can be a key to search of the conditions and factors, which lead to the production of such compounds.

Depending on biological function of the suggested molecules, the development of monitoring and the amount controlling methods could be beneficial for both medicine and biotechnology to have more advanced tools for metabolism manipulation of microorganisms and to overcome resistance against medicines.

3.3. Further improvements

Although, the developed method is already working, most steps of the it could be further optimized.

Possible good addition in the purification steps is a column, which can separate by different amount of phosphorus in the compound, as it would discriminate polyphosphates already on this stage, making the sample purer for analytical separation, as well as it would serve as better confirmation of the group of compounds, that we are dealing with.

The most crucial step to be taken in order to have confidence in the compounds' candidates, is use of the LC-MS^N system. As was reported for discovered compound, there may be peak overlaps at specified *m/z* ratio before fragmentation (Spangler *et al.* 2010; Walker and Berkowitz 2013). Therefore, use of triple quadrupole or other tandem mass spectrometer for investigation of this issue is inevitable.

As aim of the method was to preliminary identify various CDNs, we were optimizing it to be able to find each of the 10 synthetic compounds at the same time. This bring to the point, that the time program and the parameters may not be perfect for each individual compound, but rather the optimal to quick search for all of them in the biological sample. Therefore, further optimization for individual compounds may be necessary for lower LOD and better separation from possible interferents.

The most problematic of all the compound is c-di-CMP. This compound elutes significantly earlier and, therefore, it is difficult to optimize the method to include it for even semi-quantification. Separate method, based on purpose would be a better alternative.

There are several steps to be optimized, in sample preparation before liquid chromatography separation. First of all, the cell destruction method should be either changed to lysis with lytic buffer with further salts removal, or fine tuning of sonication procedure. Another point in sample preparation part is estimation of growing phases of bacteria. These steps must be taken with stronger time precision (at least by hour instead of approximate phases period, taken from the literature).

For better quantification, internal standard method could be used in addition to spiking. This would allow give high precision recovery information for the whole sample preparation procedure. In addition, establishing of high-specific method for quantification of bacteria in culture is preferable. The simplest of such methods is the protein assays, such as those after Bradford (Zheng et al. 2015) or after Smith (Bai et al. 2012). Currently, procedure after Bradford is being implemented and should be introduced before switching to tandem mass spectrometry and the part of each sample is already being taken for this assay.

CONCLUSION

Cyclic dinucleotides, although known for a long time, received considerable attention only recently. These compounds play central role in various bacterial key processes from motility to triggering pathogenesis and also play role in mammalian response to the pathogens. They are synthesized and degraded by distinct group of proteins. Dinucleotide cyclases have high specificity towards not only nucleotides, but also positions of their cyclisation. Phosphodiesterases of cyclic dinucleotides besides degrading function, often play role of the intracellular receptor.

Until now, only three compounds were found and confirmed in nature. However, many proteins with specific characteristic domains do not react on the discovered compounds *in vitro*. This fact suggests that there may be other compounds with close properties and the necessity of their discovery is undoubtable.

The main goal of this research was to develop the method, which could be suitable to identify possible candidate compounds of cyclic dinucleotides in bacteria extracts by liquid chromatography - mass spectrometry. This goal was achieved with following results.

Method development was based on the synthetic cyclic dinucleotide compounds, synthesized in-house. Sample preparation involved several purification steps and led to separation of dinucleotide fraction with low contamination by compounds with different properties and structure.

For analysis, separation of standard was set to be performed by two C18 columns with different properties. GravitySB to use for everyday measurement and Isis for confirmation of findings in biological samples. Peaks were narrow (around 0.3 minutes), but only peaks on GravitySB were symmetrical.

Method had limit of detection of 0.12 ng/ml or lower, depending on compound, with the exception of c-di-CMP, which was not possible to optimize for eluting with other compounds. Linear range on GravitySB was between 1.5 ng/ml and 15 ng/ml, which is sufficient for biological samples. Method was highly stable with standard deviation for standard compounds not exceeding 20% for GravitySB and only 9.7% at highest for Isis.

For mass spectrometry, negative mode was chosen as it gave higher ionization of the target compounds.

Analysis of biological samples revealed, that method is capable to confirm in most cases findings from literature, as well as candidates for some new compounds. It was also

shown, that some of this candidates change their abundance in growing phase-dependent manner.

To confirm the findings, it is important to perform measurements on more refined mass spectrometer with fragmentation.

ACKNOWLEDGEMENT

First and foremost, I owe my deepest gratitude to my Supervisors: Jörg Hartig, for taking me into the project, ultimate support and advices, he gave me, and Koit Herodes for helping till the very last minute in bringing to perfection my thesis distantly, as well as huge encouragement towards the development of project. In addition, I am grateful to Uno Mäeorg, my Erasmus coordinator, for his strong support and help in finding the laboratory of Prof. Hartig.

I am indebted to my many of my colleagues for their taking part in the project, in particular Changhao Wang, Dmitry Galetsky, Anna Heiler and Astrid Joachimi. Besides, I would like to show my gratitude to Patriícia Dérer, who was main corrector of my thesis paper and helped in many other ways as well.

It is an honor for me to be part of the Applied Measurement Science program and I owe deep gratitude to Ivo Leito for creating and sustaining the program, as well as to my groupmates.

BIBLIOGRAPHY

Baba T., Ara T., Hasegawa M., Takai Y., Okumura Y., Baba M., Datsenko K. A., Tomita M., Wanner B. L., Mori H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008.

Bai Y., Yang J., Zhou X., Ding X., Eisele L. E., Bai G. (2012) Mycobacterium tuberculosis Rv3586 (DacA) is a diadenylate cyclase that converts ATP or ADP into c-di-amp. *PLoS One* **7**, 1–10.

Bejerano-Sagie M., Oppenheimer-Shaanan Y., Berlatzky I., Rouvinski A., Meyerovich M., Ben-Yehuda S. (2006) A Checkpoint Protein That Scans the Chromosome for Damage at the Start of Sporulation in Bacillus subtilis. *Cell* **125**, 679–690.

Caly D., Bellini D. (2015) Targeting Cyclic di-GMP Signalling: A Strategy to Control Biofilm Formation? *Curr. Pharm. Des.* **21**, 12–24.

Florkin M. (1991) The discovery of adenosine triphosphate and the establishment of its structure. *J. Hist. Biol.* **24**, 145–154.

Francis S. H., Corbin J. D. (1999) Cyclic nucleotide-dependent protein kinases: intracellular receptors for cAMP and cGMP action. *Crit. Rev. Clin. Lab. Sci.* **36**, 275–328.

Gaffney B. L., Veliath E., Zhao J., Jones R. a (2010) One-Flask Syntheses of c-di-GMP and the [R p , R p] and [R p , S p] Thiophosphate Analogues. *Org. Lett.* **12**, 3269–3271.

Gao J., Tao J., Liang W., Zhao M., Du X., Cui S., Duan H., Kan B., Su X., Jiang Z. (2015) Identification and characterization of phosphodiesterases that specifically degrade 3′3′-cyclic GMP-AMP. *Cell Res.* **25**, 539–550.

Gao P., Ascano M., Wu Y., Barchet W., Gaffney B. L., Zillinger T., Serganov A. A., et al. (2013) Cyclic [G(2`,5`)pA(3`,5`)p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. *Cell* **153**, 1094–1107.

Hallberg Z. F., Wang X. C., Wright T. A., Nan B., Ad O., Yeo J., Hammond M. C. (2016) Hybrid promiscuous (Hypr) GGDEF enzymes produce cyclic AMP-GMP (3′, 3′-cGAMP). *Proc. Natl. Acad. Sci.* **113**, 1790–1795.

Hayakawa Y., Nagata R., Hirata A., Hyodo M., Kawai R. (2003) A facile synthesis of cyclic bis(3`5`)diguanylic acid. *Tetrahedron* **59**, 6465–6471.

He Q., Wang F., Liu S., Zhu D., Cong H., Gao F., Li B., et al. (2016) Structural and Biochemical Insight into the Mechanism of Rv2837c from Mycobacterium tuberculosis as a c-di-NMP Phosphodiesterase. *J. Biol. Chem.* **291**, 3668–81.

Huynh T. N., Woodward J. J. (2016) Too much of a good thing: regulated depletion of c-di-

AMP in the bacterial cytoplasm. *Curr. Opin. Microbiol.* **30**, 22–29.

Hyodo M., Hayakawa Y. (2008) Synthesis, Chemical Properties and Biological Activities of Cyclic Bis(3′–5′)diguanylic Acid (c-di-GMP) and its Analogues, in *Modif. Nucleosides*, (Herdewijn P., ed), pp. 343–363. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.

Kalia D., Merey G., Nakayama S., Zheng Y., Zhou J., Luo Y., Guo M., Roembke B. T., Sintim H. O. (2013) Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria and implications in pathogenesis. *Chem Soc Rev* **42**, 305–341.

Kelsey I., Nakayama S., Sintim H. O. (2012) Diamidinium and iminium aromatics as new aggregators of the bacterial signaling molecule, c-di-GMP. *Bioorganic Med. Chem. Lett.* **22**, 881–885.

Kolb A., Busby S., Buc H., Garges S., Adhya S. (1993) Transcriptional regulation by cAMP and its receptors. *Ann.Rev.Biochem.* **62**, 749–795.

Krasteva P. V., Giglio K. M., Sondermann H. (2012) Sensing the messenger: The diverse ways that bacteria signal through c-di-GMP. *Protein Sci.* **21**, 929–948.

Lehmann A. (2003) Watson and DNA: Making a Scientific Revolution. *DNA Repair (Amst).* **2**, 1073–1074.

MACHEREY-NAGEL GmbH & Co. KG (2015) *C18 Gravity-SB*. Düren, Germany.

MACHEREY-NAGEL GmbH & Co. KG. (2015) *Nucleodur Isis.pdf*. Düren.

Nakayama S., Kelsey I., Wang J., Roelofs K., Stefane B., Luo Y., Lee V. T., Sintim H. O. (2011) Thiazole orange-induced c-di-GMP quadruplex formation facilitates a simple fluorescent detection of this ubiquitous biofilm regulating molecule. *J. Am. Chem. Soc.* **133**, 4856–4864.

Nakayama S., Roelofs K., Lee V. T., Sintim H. O. (2012) A C-di-GMP–proflavine–hemin supramolecular complex has peroxidase activity—implication for a simple colorimetric detection. *Mol. Biosyst.* **8**, 726.

National Center for Biotechnology Information (2016) *c-di-GMP*.

Paul K., Nieto V., Carlquist W. C., Blair D. F., Harshey R. M. (2010) The c-di-GMP Binding Protein YcgR Controls Flagellar Motor Direction and Speed to Affect Chemotaxis by a "Backstop Brake" Mechanism. *Mol. Cell* **38**, 128–139.

Potrykus K., Cashel M. (2008) (p)ppGpp: still magical? *Annu. Rev. Microbiol.* **62**, 35–51.

Romling U. (2008) Great times for small molecules: c-di-AMP, a second messenger candidate in Bacteria and Archaea. *Sci Signal* **1**, pe39.

Römling U. (2002) Molecular biology of cellulose production in bacteria. *Res. Microbiol.*

153, 205–212.

Romling U., Galperin M. Y., Gomelsky M. (2013) Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.* **77**, 1–52.

Ross P., Weinhouse H., Aloni Y., Michaeli D., Weinberger-Ohana P., Mayer R., Braun S., et al. (1987) *Regulation of cellulose synthesis in Acetobacter xylinum by cyclic diguanylic acid*.

Schirmer T., Jenal U. (2009) Structural and mechanistic determinants of c-di-GMP signalling. *Nat. Rev. Microbiol.* **7**, 724–735.

Spangler C., Böhm A., Jenal U., Seifert R., Kaever V. (2010) A liquid chromatography-coupled tandem mass spectrometry method for quantitation of cyclic di-guanosine monophosphate. *J. Microbiol. Methods* **81**, 226–231.

Sudarsan N., Lee E. R., Weinberg Z., Moy R. H., Kim J. N., Link K. H., Breaker R. R. (2008) Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* **321**, 411–3.

Sun L., Wu J., Du F., Chen X., Chen Z. J. (2013) Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786–91.

Sutherland E. W., Robison G. A., Butcher W. R. (1968) Some Aspects of the Biological Role of Adenosine 3',5'-monophosphate (Cyclic AMP). *Circulation* **37**, 279–306.

Tischler A. D., Camilli A. (2005) Cyclic diguanylate regulates Vibrio cholerae virulence gene expression. *Infect. Immun.* **73**, 5873–82.

Tschowri N. (2016) Cyclic dinucleotide-controlled regulatory pathways in Streptomyces species. *J. Bacteriol.* **198**, 47–54.

Walker R. K., Berkowitz G. A. (2013) Cyclic Nucleotide Signaling in Plants. *Cycl. Nucleotide Signal. Plants Methods Mol. Biol.* **1016**, 245–252.

Witte G., Hartung S., Büttner K., Hopfner K. P. (2008) Structural Biochemistry of a Bacterial Checkpoint Protein Reveals Diadenylate Cyclase Activity Regulated by DNA Recombination Intermediates. *Mol. Cell* **30**, 167–178.

Wu J., Sun L., Chen X., Du F., Shi H., Chen C., Chen Z. J. (2013) Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science* **339**, 826–30.

Zhang Z., Kim S., Gaffney B. L., Jones R. A. (2006) Polymorphism of the signaling molecule c-di-GMP. *J. Am. Chem. Soc.* **128**, 7015–7024.

Zheng C., Ma Y., Wang X., Xie Y., Ali M. K., He J. (2015) Functional analysis of the sporulation-specific diadenylate cyclase CdaS in Bacillus thuringiensis. **6**, 1–15.

Zhou H., Zheng C., Su J., Chen B., Fu Y., Xie Y., Tang Q., Chou S.-H., He J. (2016) Characterization of a natural triple-tandem c-di-GMP riboswitch and application of the riboswitch-based dual-fluorescence reporter. *Sci. Rep.* **6**, 20871.

Zogaj X., Nimtz M., Rohde M., Bokranz W., Römling U. (2001) The multicellular morphotypes of Salmonella typhimurium and Escherichia coli produce cellulose as the second component of the extracellular matrix. *Mol. Microbiol.* **39**, 1452–1463.

ANNEXES

Annex A. Media composition

LB medium

1. Add the following to 800.0 ml H_2O
   - Tryptone. 10.0 g
   - Yeast extract. 5.0 g
   - NaCl. 10.0 g
2. Adjust pH to 7.5 with NaOH.
3. Adjust volume to 1.0 l with dH_2O
4. Sterilize by autoclaving

M9 medium

1. Make M9 Salts:
   - aliquot 800.0 ml H_2O
   - $Na_2HPO_4 \times 7\ H_2O$ 64.0 g
   - KH_2PO_4 15.0 g
   - NaCl 2.5 g
   - NH_4Cl 5.0 g
   - Stir until dissolved
   - Adjust to 1000ml with distilled H2O
   - Sterilize by autoclaving
2. Measure ~700.0 ml of distilled H_2O (sterile)
3. Add 200.0 ml of M9 salts
4. Add 2.0 ml of 1 M $MgSO_4$ (sterile)
5. Add 20.0 ml of 20% glucose (or other carbon source)
6. Add 100.0 µl of 1 M $CaCl_2$ (sterile)
7. Adjust to 1000 ml with distilled H_2O

MSgg medium

1. 50.0 µM $MnCl_2$
2. 5.0 mM KH_2PO_4
3. 1.0 µM $ZnCl_2$
4. 50.0 µM $FeCl_3$
5. 2.0 mM $MgCl_2$
6. 700.0 µM $CaCl_2$
7. 50.0 µg/ml threonine, tryptophan, and phenylalanine
8. 0.5% glutamate
9. 0.5% glycerol
10. 2.0 µM thiamine
11. 100.0 mM morpholinepropanesulfonic acid (MOPS) (pH 7)

GYM Streptomyces medium (medium 65)

1. Glucose 4.0 g
2. Yeast extract 4.0 g
3. Malt extract 10.0 g
4. $CaCO_3$ 2.0 g
5. Agar 12.0 g
6. Distilled water 1000.0 ml
7. Adjust pH to 7.2 before adding agar. Delete $CaCO_3$ if liquid medium is used.

PYE medium

1. K_2HPO_4 1.00 g
2. $MgSO_4$ x 7 H_2O 0.20 g
3. CaCl x 2 H_2O 0.02 g
4. $Na_2S_2O_3$ x 5 H_2O 0.20 g
5. Na-pyruvate 2.20 g
6. Yeast extract 4.00 g
7. $(NH_4)_2SO_4$ 1.00 g
8. Trace element solution SL6 (see Medium 27) 1.00 ml Distilled water
9. 1000.00 ml
10. Adjust pH to 6.8.
11. Boil the medium for few minutes. Fill in tubes with rubber stoppers under nitrogen gas. Sterilize at 121°C for 15 min.

Glucose yeast extract medium (medium 54)

1. Glucose 20.0 g
2. Yeast extract 10.0 g
3. $CaCO_3$(light precipitate) 20.0 g
4. Distilled water 1000.0 ml
5. For solid medium, add 17.0 g/l agar.

TY medium

1. H_2O 1 L
2. Tryptone 6.0 g

3. Yeast extract 3.0 g
4. CaCl2×2 H_2O 0.5 g
5. Agar-agar (optional) 12.0 g
6. agar 3.0 g + broth per bottle 250.0 ml

YPAD medium

1. Yeast extract 1% 10.0 g
2. Peptone 2% 20.0 g
3. Glucose 2% 20.0 g
4. Agar 2% 20.0 g
5. Adenine sulfate 0.004% 40.0mg

Annex B. Limit of detection

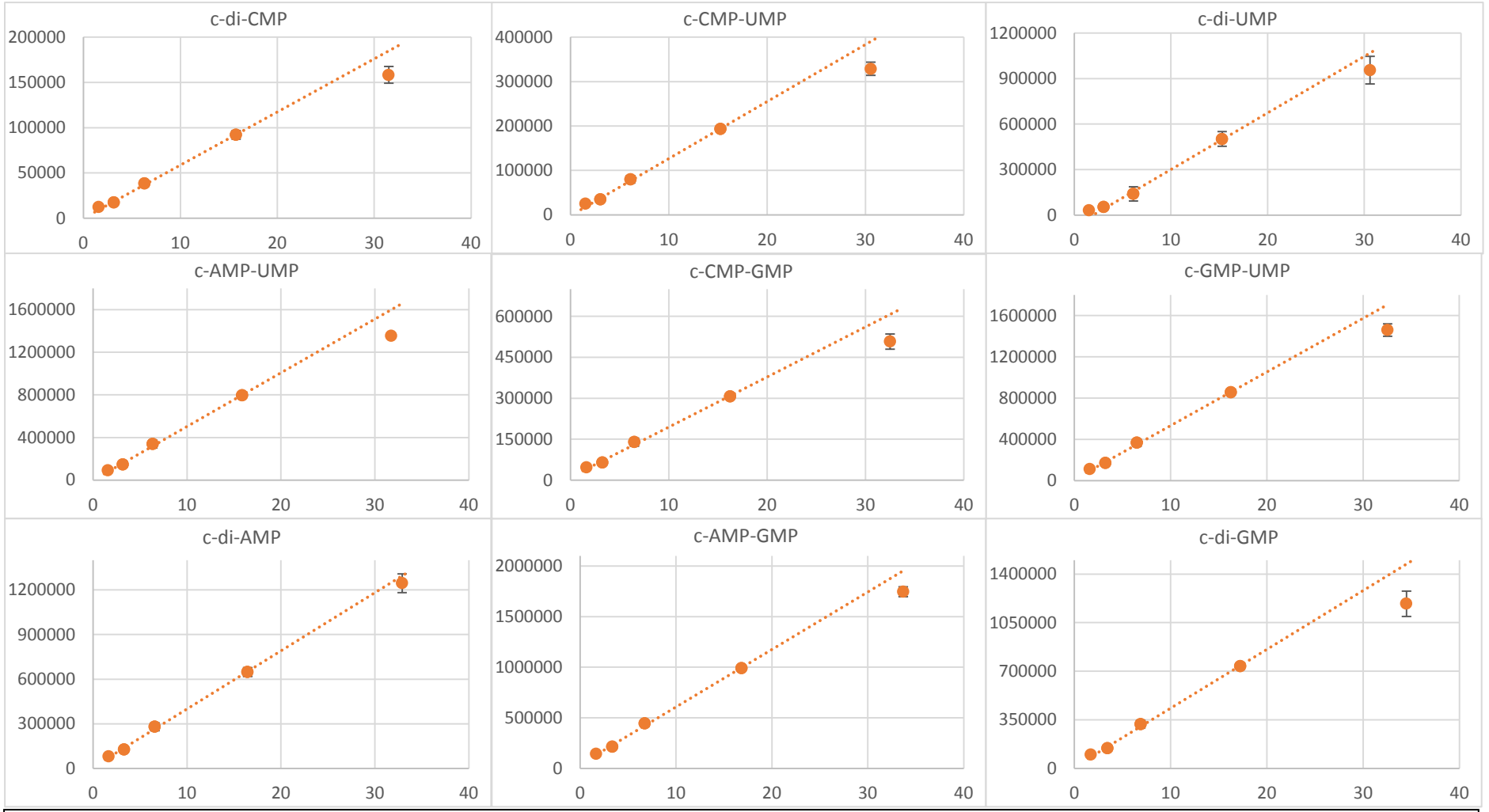
C18 Isis

E. coli										Average	St Dev	LOD (µV)	Standards	M_r, Da	LOD (ng/ml)	CDN
714	539	515	535	547	518	580	637	549	538	567.2	62.50	773.45	4564.82	610	**1.034**	c-di-CMP
277	403	380	350	363	354	370	353	370	320	354	34.60	468.17	59514.82	611	**0.048**	c-CUMP
415	323	320	293	330	315	343	320	312	333	330.4	32.61	438.00	52133.82	612	**0.051**	c-di-UMP
293	341	369	328	369	332	399	414	440	380	366.5	44.26	512.57	130293.36	634	**0.025**	c-CAMP
270	334	324	335	324	357	381	395	378	428	352.6	44.81	500.46	168624.18	635	**0.019**	c-UAMP
436	279	463	323	395	378	359	415	338	322	370.8	57.40	560.22	117074.18	650	**0.031**	c-CGMP
253	384	415	420	392	419	396	437	389	358	386.3	51.97	557.80	253986.55	651	**0.014**	c-UGMP
637	154	202	173	193	210	199	228	237	235	246.8	139.62	707.55	198145.64	658	**0.023**	c-di-AMP
634	986	860	1052	974	958	1013	960	993	1051	948.1	123.10	1354.32	353129.55	674	**0.026**	c-AGMP
216	583	469	444	467	436	465	434	486	425	442.5	91.39	744.08	222148.91	690	**0.023**	c-di-GMP

C18 GravitySB

E. coli										Average	St Dev	LOD (µV)	Standards	M_r, Da	LOD (ng/ml)	CDN
433	247	446	428	284	472	499	546	463	514	433.2	96.09	750.29	20444.73	610	**0.224**	c-di-CMP
544	437	392	333	388	457	363	441	499	354	420.8	67.24	642.69	44774.48	611	**0.088**	c-CUMP
540	799	655	378	803	613	701	466	878	424	625.7	172.14	1193.77	64785.68	612	**0.113**	c-di-UMP
476	447	306	299	320	401	306	438	476	514	398.3	83.30	673.18	119151.71	634	**0.036**	c-CAMP
494	332	430	462	463	425	418	500	517	416	445.7	53.94	623.69	171394.13	635	**0.023**	c-UAMP
605	278	506	334	456	407	419	399	366	394	416.4	91.01	716.74	72385.50	650	**0.064**	c-CGMP
436	445	686	369	634	469	349	521	486	525	492.0	106.05	841.96	205552.77	651	**0.027**	c-UGMP
705	556	376	553	407	539	584	396	519	524	515.9	99.89	845.53	152720.92	658	**0.036**	c-di-AMP
394	656	575	279	647	557	371	561	625	480	514.5	128.75	939.36	262913.65	674	**0.024**	c-AGMP
731	282	543	452	664	292	624	370	334	621	491.3	166.30	1040.10	180203.77	690	**0.040**	c-di-GMP

Annex C. Calibration curves prepared on C18 GravitySB column.



For all the graphs: values on abscissa – concentration in ng/ml; on ordinate – intensity in µV

Annex D Standards' reproducibility

C18 Isis

c-di-CMP	c-CUMP	c-di-UMP	c-CAMP	c-UAMP	c-CGMP	c-UGMP	c-di-AMP	c-AGMP	c-di-GMP	Compounds
4580	59378	50753	132299	179749	108432	255837	189029	381151	238091	**22.May**
4250	62928	52138	134006	175406	131394	249317	188369	361087	218071	**19.May**
4858	60645	48397	141439	156945	125513	238376	192581	349766	217623	**18.May**
4360	58442	46065	135664	172616	115278	261342	217580	326578	254856	**17.May**
4608	62554	56831	123148	178986	110381	267603	201466	351193	218408	**14.May**
4036	57192	53789	130611	165247	117928	259576	175687	332860	227048	**06.May**
4525	63044	55961	135938	177915	117013	246282	207932	372289	211812	**30.April**
5526	60800	55656	135937	166097	114873	250844	190712	312135	207409	**21.April**
5047	53371	53314	124756	139102	112287	247274	194647	366600	208061	**19.April**
4343	53236	45158	129576	165295	113876	272118	224655	354608	197518	**13.April**
4080	63073	55410	109853	177508	120841	245283	196944	376158	244741	**06.April**
4564.818	**59514.82**	**52133.82**	**130293.4**	**168624.2**	**117074.2**	**253986.5**	**198145.6**	**353129.5**	**222148.9**	**average**
441.6976	**3642.34**	**4066.27**	**8571.641**	**12227.58**	**6722.317**	**10288.18**	**14036.5**	**21773.86**	**17431.82**	**standard deviation**
9.68%	**6.12%**	**7.80%**	**6.58%**	**7.25%**	**5.74%**	**4.05%**	**7.08%**	**6.17%**	**7.85%**	**%stdev**
10.92%	**6.22%**	**9.18%**	**16.97%**	**1.25%**	**11.44%**	**4.13%**	**4.19%**	**1.31%**	**2.79%**	**%first/last**

C18 GravitSB

c-di-CMP	c-CUMP	c-di-UMP	c-CAMP	c-UAMP	c-CGMP	c-UGMP	c-di-AMP	c-AGMP	c-di-GMP	Compounds
17804	50307	65953	112653	185304	81073	194335	142730	254600	218428	**23.May**
22444	45443	63884	109613	137961	74475	178219	145356	212154	169770	**05.May**
18162	47464	56848	113671	164417	70695	229873	158476	222075	176214	**04.May**
23975	56494	41672	116593	183533	70032	174711	147211	243843	193235	**03.May**
25038	41165	62797	115359	166768	56314	216391	129186	248475	173264	**02.May**
21518	34674	73491	119724	179453	76767	226730	155478	291387	197939	**28.April**
23228	44288	54479	123450	186477	75505	212878	134947	241905	167943	**27.April**

c-di-CMP	c-CUMP	c-di-UMP	c-CAMP	c-UAMP	c-CGMP	c-UGMP	c-di-AMP	c-AGMP	c-di-GMP	Compounds
19763	47733	65733	148911	187081	79583	188103	153099	268910	205763	**26.April**
17147	50156	78429	128115	177889	67609	220415	176763	310916	203593	**20.April**
23300	50110	50415	117835	205124	84359	209263	142322	295580	179777	**18.April**
20168	44010	56642	112234	158076	73382	209261	161525	288388	190354	**14.April**
20205	42845	74740	120512	178478	71923	227284	140482	246343	164743	**11.April**
16556	45076	71880	113593	164198	75310	216974	151266	231543	176155	**08.April**
22521	36014	53214	110940	153725	73508	200755	156339	262596	140423	**07.April**
16771	48919	58912	127301	157099	69574	212152	150398	268966	154221	**06.April**
19249	38879	81113	116809	208553	76970	217919	138443	310301	157007	**05.April**
22324	50659	68311	120011	139382	61283	191406	149601	333441	175935	**04.April**
26161	39498	78891	140322	165608	64052	241905	156333	273174	138682	**01.April**
17342	48948	95996	144522	171971	71431	134592	251181	220250	145441	**30.March**
19387	46990	60643	113567	177848	65654	213054	164903	281125	207222	**16.March**
20055	43893	55945	117940	165583	71322	208930	152976	266137	191791	**14.March**
23800	50195	58835	125495	180004	71870	225409	163509	292744	210587	**11.March**
20836	51204	60066	118906	176971	68824	216934	159544	283698	206634	**08.March**
17544	37597	57351	104713	154556	62329	182311	139311	233048	169694	**07.March**
17947	38204	68918	103502	178430	76926	204180	156297	261256	186693	**04.March**
18110	38618	61296	116633	171246	80040	209033	161740	275334	199762	**02.March**
23371	55650	66521	132362	186213	80461	228034	165235	291597	207258	**01.March**
21837	44872	96201	101280	196627	76078	209948	146743	238007	171731	**26.February**
20950	43462	63324	125457	170804	83206	222886	150259	272400	177643	**28.January**
15274	31917	41515	82211	121186	57603	155401	106571	198155	132585	**26.January**
20999.5	42725	64341	139469	162653	75792.5	192850	126124.5	231975	195830	**22.January**
20444.73	**44774.48**	**64785.68**	**119151.7**	**171394.1**	**72385.5**	**205552.8**	**152720.9**	**262913.6**	**180203.8**	**average**
2763.133	**5968.215**	**12631.26**	**13325.36**	**18663.66**	**7054.084**	**22814.3**	**22782.4**	**31639.27**	**23034.14**	**standard deviation**
13.52%	**13.33%**	**19.50%**	**11.18%**	**10.89%**	**9.75%**	**11.10%**	**14.92%**	**12.03%**	**12.78%**	**%stdev**
15.22%	**17.75%**	**2.51%**	**19.23%**	**13.93%**	**6.97%**	**0.77%**	**13.17%**	**9.75%**	**11.54%**	**%first/last**

Annex E. Measurement results for different columns

C18 Gravity SB

Sample information						Cyclic dinucleotides									
Code	**Organism**	**ssp.**	**Medium**	**Phase**	**Date of first culture[2]**	**c-di-CMP**	**c-CUMP**	**c-di-UMP**	**c-CAMP**	**c-UAMP**	**c-CGMP**	**c-UGMP**	**c-di-AMP**	**c-AGMP**	**c-di-GMP**
#15a	*St. cohnii*	GH137	LB	exponential	02.Mar			**0.45**	**1.02**			**0.25**			
#15b	*St. cohnii*	GH137	LB	stationary	02.Mar				**0.97**			**0.47**			
#17a	*C. testosteroni*	KF-1	LB	exponential	02.Mar		**0.46**		**0.43**		**0.62**		**1.54**		**0.27**
#17b	*C. testosteroni*	KF-1	LB	stationary	02.Mar	**0.13**	**0.19**		**0.14**	**0.07**	**0.09**	**0.13**	**0.17**	**0.09**	
#18a	*Ps. putida*	KT-2440	LB	exponential	02.Mar							**1.71**			
#18b	*Ps. putida*	KT-2440	LB	stationary	02.Mar					**0.03**		**0.05**		**0.04**	**0.51**
#4b	*B. subtilis*	DSM10	LB	exponential	02.Mar	**2.34**			**2.77**						
#4c	*B. subtilis*	DSM10	LB	stationary	02.Mar	**1.19**									
#4d	*B. subtilis*	DSM10	MSgg	biofilm	02.Mar	**4.52**		**2.80**	**0.46**				**0.91**		
#14a	*S. erythraea*	ATCC11635	Medium 65	exponential	03.Mar			**11.48**			**1.26**				**0.99**
#14b	*S. erythraea*	ATCC11635	Medium 65	stationary	03.Mar										**0.06**
#16a	*S. cerevisiae*	BY4741	YPAD	exponential	12.Apr			**0.17**	**1.97**	**0.33**		**0.34**	**0.35**	**0.34**	**0.41**
#16b	*S. cerevisiae*	BY4741	YPAD	stationary	12.Apr										

[2] Second preparation was strictli in 14working days after the first

Sample information						Cyclic dinucleotides									
Code	**Organism**	**ssp.**	**Medium**	**Phase**	**Date of first culture[2]**	**c-di-CMP**	**c-CUMP**	**c-di-UMP**	**c-CAMP**	**c-UAMP**	**c-CGMP**	**c-UGMP**	**c-di-AMP**	**c-AGMP**	**c-di-GMP**
#19a	*C. crescentus*	CB15N (NA1000)	PYE	exponential	15.Apr			**4.84**	**1.02**						**1.24**
#19b	*C. crescentus*	CB15N (NA1000)	PYE	stationary	15.Apr	**0.52**		**3.41**	**0.30**	**0.19**	**0.19**	**0.21**	**0.34**	**0.16**	**0.28**
#20a[3]	*X. campestris*	ATCC 33913	Medium 54,	stationary	10.May		**2.84**	**1.71**	**1.41**	**0.63**	**3.80**	**1.29**	**0.88**	**0.70**	**0.80**
#20b	*X. campestris*	ATCC 33913	Medium 54,	stationary	10.May				**0.12**		**0.60**	**2.90**	**0.12**	**2.40**	**0.25**
#20c	*X. campestris*	ATCC 33913	TY	stationary	11.May				**1.28**			**3.45**	**1.41**	**0.49**	**0.34**
#20d	*X. campestris*	ATCC 33913	TY	stationary	11.May							**3.17**	**2.30**	**0.75**	**0.26**

[3] *X. campestris* is measured only once so far

C18 Isis

Sample information						Cyclic dinucleotides									
Code	**Organism**	**ssp.**	**Medium**	**Phase**	**Date of first culture**	**c-di-CMP**	**c-CUMP**	**c-di-UMP**	**c-CAMP**	**c-UAMP**	**c-CGMP**	**c-UGMP**	**c-di-AMP**	**c-AGMP**	**c-di-GMP**
#15a	*St. cohnii*	GH137	LB	exponential	02.Mar			**yes**	**yes**						
#15b	*St. cohnii*	GH137	LB	stationary	02.Mar		**yes**		**yes**						
#17a	*C. testosteroni*	KF-1	LB	exponential	02.Mar		**yes**		**yes**				**yes**		**yes**
#17b	*C. testosteroni*	KF-1	LB	stationary	02.Mar										**yes**
#18a	*Ps. putida*	KT-2440	LB	exponential	02.Mar							**yes**			
#18b	*Ps. putida*	KT-2440	LB	stationary	02.Mar										**yes**
#4b	*B. subtilis*	DSM10	LB	exponential	02.Mar				**yes**						
#4c	*B. subtilis*	DSM10	LB	stationary	02.Mar	**yes**					**yes**				
#4d	*B. subtilis*	DSM10	MSgg	biofilm	02.Mar	**yes**		**yes**						**yes**	
#14a	*S. erythraea*	ATCC11635	Medium 65	exponential	03.Mar			**yes**			**yes**				**yes**
#14b	*S. erythraea*	ATCC11635	Medium 65	stationary	03.Mar								**yes**		
#16a	*S. cerevisiae*	BY4741	YPAD	exponential	12.Apr		**yes**		**yes**						**yes**
#16b	*S. cerevisiae*	BY4741	YPAD	stationary	12.Apr										
#19a	*C. crescentus*	CB15N (NA1000)	PYE	exponential	15.Apr			**yes**	**yes**						**yes**
#19b	*C. crescentus*	CB15N (NA1000)	PYE	stationary	15.Apr			**yes**							

Sample information						Cyclic dinucleotides									
Code	**Organism**	**ssp.**	**Medium**	**Phase**	**Date of first culture**	**c-di-CMP**	**c-CUMP**	**c-di-UMP**	**c-CAMP**	**c-UAMP**	**c-CGMP**	**c-UGMP**	**c-di-AMP**	**c-AGMP**	**c-di-GMP**
#20a	*X. campestris*	ATCC 33913	Medium 54,	stationary	10.May		**yes**		**yes**	**yes**	**yes**	**yes**	**yes**		**yes**
#20b	*X. campestris*	ATCC 33913	Medium 54,	stationary	10.May		**yes**	**yes**			**yes**	**yes**		**yes**	
#20c	*X. campestris*	ATCC 33913	TY	stationary	11.May	**yes**			**yes**			**yes**	**yes**	**yes**	
#20d	*X. campestris*	ATCC 33913	TY	stationary	11.May				**yes**			**yes**	**yes**	**yes**	

INFORMATION SHEET

Pealkiri: Bakterikultuurides tsükliliste dinukleotiidide esmase identifitseerimismetoodika arendamine.

Kokkuvõte: Tsüklilised dinukleotiidid on eeldatavasti teisased virgatsained bakterites. Mõned selle aineklassi teadaolevad ühendid on olulisel kohal paljudes rakus toimuvates protsessides alates raku elutsüklist kuni patogeneesini. *In vivo* ühendeid sünteesitakse tsükliliste dinukleotiidi tsüklaaside ja lagundatakse spetsiifiliste fosfodiesteraaside vahendusel. Paljud teadaolevad sellesse klassi kuuluvad valgud ei interakteeru tuntud tsükliliste dinukleotiididega, mis viitab sellele, et rakus võib esineda senitundmatuid tsüklilisi dinukleotiide. Käesolevas töös töötati välja mitut kolonni kasutav vedelikkromatograafia-massispektromeetriline metoodika huvipakkuvate ainete määramiseks bakteriekstraktist. Metoodika aluseks on sünteesitud tsükliliste nukleotiidide standardained. Väljatöötatud analüüsimetoodika avastamispiir on madalam, kui teaduskirjanduses siiani avaldatud üksikute tsükliliste dinukleotiidide metoodikatel. Metoodika toimivuse kinnituseks uuriti mitmeid bakterikultuure, mis teadaolevalt produtseerivad diguanülaadi või diadenülaadi tsüklilisi dinukleotiide. Samuti näidati, et mõnede tsükliliste dinukleotiidide tase sõltub bakterite kasvufaasist.

Märksõnad: bakterid: ainevahetus, tsüklilised dinukleotiidid, c-di-GMP, HPLC, LC-MS, kvalitatiivne analüüs, teisane virgatsaine.

Title: Development of method for preliminary identification of cyclic dinucleotides in bacterial cultures

Summary: Cyclic dinucleotides are perspective second messengers in *Bacteria*. Known compounds of the class play various roles in many processes in bacteria from cell cycle to pathogenesis. *In vivo* compounds are synthesized with cyclic dinucleotide cyclases and degrade with specific phosphodiesterases. Many of discovered proteins of these classes do not interact with known cyclic dinucleotide compounds, which suggest another compounds of the class may be also in the bacterial cell. In this work a method was developed , based on liquid chromatography coupled mass spectrometry with several columns, which is capable to distinguish possible compound of interest in the bacterial extract based on the chemically synthesized compounds of the class. Method is shown to have LOD lower, than in any method proposed before for single cyclic dinucleotide compound. As a proof of principle, several bacteria reported to produce cyclic diguanylate or cyclic diadenylate were tested. In addition, it was demonstrated that levels of some of cyclic dinucleotides are dependent on growing phase of bacteria.

Keywords: Bacteria: metabolism, cyclic dinucleotides, c-di-GMP, HPLC, LC-MS, qualitative analysis, second messenger.

CERCS code: P300, P320

Non-exclusive licence to reproduce thesis and make thesis public

I, Stanislav Andres (02.06.1993),

(*author's name*)

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

1.2. make available to the public via the university's web environment, including via the DSpace digital archives, as of **01.09.2017** until expiry of the term of validity of the copyright,

"Development of method for preliminary identification of cyclic dinucleotides in bacterial cultures.",

supervised by Prof. Dr. Jörg Hartig (University of Konstanz) and Associate professor Koit Herodes (University of Tartu),

2. I am aware of the fact that the author retains these rights.

3. This is to certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Konstanz, **30.05.2016**