Evolutional and functional diversification of HD-GYP phosphodiesterases; metal content and substrate specificity

Master’s Thesis

Presented to

The Faculty of the Graduate School of Arts and Sciences
Brandeis University
Department of Biochemistry
Maria-Eirini Pandelia, Advisor

In Partial Fulfillment
of the Requirements for
Master of Science
in
Biochemistry and Biophysics

by
Sining Sun
May 2018
Abstract

Evolutional and functional diversification of HD-GYP phosphodiesterases; metal content and substrate specificity

A thesis presented to the Department of Biochemistry

Graduate School of Arts and Sciences
Brandeis University
Waltham, Massachusetts

By Sining Sun

Cyclic dinucleotides (CDNs) act as intracellular messengers and modulate many cellular activities including innate immune activation, virulence factor production and biofilm formation in bacterial pathogens. The cellular levels of CDNs are fine-tuned by cyclases (involved in their synthesis) and phosphodiesterases (PDEs) (involved in their degradation). HD-GYP proteins represent a relative novel addition to the PDE functional superfamily, and belong to a large superfamily of proteins annotated as HD-domain, designated by the tandem histidine-aspartate as a core structural motif. In this study, we aimed to determine the metal cofactor specificity of HD-GYP PDEs and catalytic proficiencies regarding different substrates and their abilities to perform one-step (to 5’-pNpN) or two-step hydrolysis (to two molecules of NMP). We initially focused on
the HD-GYP protein from Vibrio cholera (VCA0681), for which no three-dimensional structure has been resolved and which is the key PDE for regulating canonical c-di-GMP and the hybrid c-GAMP. VCA0681 exhibits a somewhat stringent Fe-dependent activity and contains binding motifs for the ligation of two diiron cofactors, with seven conserved protein residues typified as HD-[HD-GYP] (the first dimetal domain lacks the GYP triad, suggested to be important for c-di-GMP hydrolysis). It was the first PDE reported to degrade the novel cyclic dinucleotide c-GAMP, associated with innate immune response. To better understand the substrate and metal specificity of this and other homologous HD-GYP PDEs, apart from exhaustive biochemical and biophysical studies, the three-dimensional structure is much needed. Though VCA0681 is considered as a prototypical PDE, it can only be isolated as a linear fusion with the maltose binding protein (MBP), removal of which leads to protein instability and precipitation, precluding any biophysical and structural investigations on the native protein. We thus carried out a bioinformatics analysis to identify a structurally and functionally homologous HD-[HD-GYP] domain protein. SO3491 from the model γ-proteobacterium Shewanella oneidensis was selected as the best candidate both due to its sequence similarity (60%) and identity (43%) to VCA0681 as well as its predictability for crystallization. I characterized the kinetic and specificity profile for SO3491 including its metal dependence and substrate specificity and demonstrated that it is a true structural and functional homolog of VCA0681. SO3491 can also hydrolyze the hybrid substrate c-GAMP. Our studies redefined the role of the second metal-binding domain, which was previously unknown. We confirmed that the C-terminal HD-GYP domain is pivotal for PDE activity. Intriguingly, the N-terminal HD-domain is essential for stopping cyclic dinucleotide degradation after the first step hydrolysis, especially in the case of the hybrid c-GAMP. To gain better insight into the HD-GYP protein superfamily, we performed a phylogenetic analysis to identify whether we could establish
a correlation between metal content, substrate specificity and mode of action for these HD-GYPs, with the ultimate aim to delineate their multiple occurrences and specific roles within the cell. This analysis reveals a biochemical and structural diversion of these proteins and provides a good predictive map for assigning metal-dependence, level of activity and reaction outcome in different HD-GYPs, allowing for constructing an evolitional diversification profile.
# Table of Contents

Abstract ................................................................................................................................. iii

List of Figures ........................................................................................................................... viii

List of Tables .......................................................................................................................... x

1. Introduction .......................................................................................................................... 1
   1.1 Cyclic dinucleotides (CDNs) ......................................................................................... 1
   1.2 HD-GYP phosphodiesterases (PDEs) ............................................................................. 3
   1.3 The *Vibrio cholera* VCA0681 and *Shewanella oneidensis* SO3491 .................. 5
   1.4 Goal of the thesis ............................................................................................................ 7

2. Experimental Procedures .................................................................................................. 7
   2.1 Material and methods .................................................................................................... 7
   2.2 Plasmid construction, cloning and mutagenesis ......................................................... 7
   2.3 Protein expression and purification .............................................................................. 9
   2.4 Cleavage of the SUMO tag by the Ulp1 protease ...................................................... 10
   2.5 Cleavage of the MBP-tag with the S219V TEV protease ......................................... 11
   2.6 Ferrozine assay ........................................................................................................... 12
   2.7 High Performance liquid chromatography (HPLC) ................................................ 12
   2.8 Multiple turnover assays .............................................................................................. 13
   2.9 Determination of steady-state of SO3491 with c-di-GMP and c-GAMP ..... 14
3. Results ........................................................................................................................................... 18

3.1 Wildtype VCA0681 and SO3491 ......................................................................................... 19

Redox State Differentiation ........................................................................................................ 24

Metal Cofactor Specificity .......................................................................................................... 27

Substrate Specificity .................................................................................................................. 29

Steady-State kinetic parameters for c-di-GMP and c-GAMP .................................................. 31

O₂-sensitivity of the ‘active’ Fe²⁺Fe²⁺ form of the wt MBP-SO3491 ........................................... 32

3.2 Individual HD-domains ......................................................................................................... 34

3.3 Bioinformatic and phylogenetic analysis .............................................................................. 40

4. Discussion .................................................................................................................................. 44

5. Future Directions and Experiments ...................................................................................... 48

Bibliography ................................................................................................................................ 49
List of Figures

Figure 1.1. Structure of cyclic dinucleotides. ................................................................. 2

Figure 1.2. Scheme of c-di-GMP hydrolysis by HD-GYP phosphodiesterases. ................. 3

Figure 1.3. Crystal structures of HD-GYP protein active sites........................................... 4

Figure 1.4. A homology model of VCA0681 featuring both HD dinuclear centers................. 6

Figure 3.1 Mössbauer spectra of wt MBP-VCA0681 and wt MBP-SO3491 ......................... 20

Figure 3.2. Mössbauer spectra of the MBP-SO3491 reduced with sodium ascorbate and sodium dithionite .................................................................................................................. 22

Figure 3.3. EPR spectra of the different forms of SO3491 and MBP-VCA0681 ...................... 23

Figure 3.4. Activity assays with ascorbate and dithionite reduced MBP-SO3491 .................. 25

Figure 3.5. Activity assays with wt MBP-SO3491 and wt MBP-VCA0681 ......................... 27

Figure 3.6. Activity assays with Mn-loaded wild-type MBP-SO3491 .................................. 28

Figure 3.7. Activity assays of wt MBP-SO3491 with different substrates ......................... 30

Figure 3.8. Steady-state kinetic parameters of SO3491 ..................................................... 31

Figure 3.9. Stop-flow experiment set up ........................................................................... 33

Figure 3.10. Stop-flow UV-VIS spectra ........................................................................... 34

Figure 3.11. Thermoshift denaturation curves of the different SO3491 constructs ............. 36
Figure 3.12. HPLC chromatograms of the activity assays with wt, D75A MBP-VCA0681 and wt, D69A MBP-SO3491 ...................................................................................................................... 37

Figure 3.13. HPLC chromatograms of the activity assays with D75A MBP-VCA0681 and c-GAMP........................................................................................................................................ 38

Figure 3.14. UV-Visible spectra of the different wt and variant SO3491 constructs in their aerobically isolated state........................................................................................................................................ 39

Figure 3.15. Maximum-likelihood rooted phylogenetic tree ................................................................................................................. 41

Figure 3.16. Unrooted phylogenetic tree ................................................................................................................................................. 41

Figure 3.17. Sequence alignment of PA4781 and tri-nuclear HD-GYP Proteins.......................................................................................... 43
List of Tables

Table 1. List of plasmid and primers ................................................................. 17

Table 2. Protein samples and metal content........................................................ 17

Table 3. Kinetic parameters of cyclic dinucleotide specific phosphodiesterases belonging to the HD-GYP and the EAL subtype. ................................................................. 32
Chapter 1

Introduction

Cyclic dinucleotides (CDNs)

Cyclic dinucleotides (CDNs) act as intracellular messengers and modulate many cellular processes, including innate immune activation, responses to small molecules (i.e. oxygen, nitric oxide), virulence factor production, cell motility, and biofilm formation in bacterial pathogens (1-3). Among all CDNs, c-di-GMP is the most intensively studied and best-understood member of the cyclic dinucleotide family of second messengers. It is an important bacterial signaling molecule involved in virulence response and biofilm formation. The biofilm formation is linked with several chronic bacterial infections, such as bloodstream infections and urinary tract infections (4, 5). Some recent studies have revealed that biofilm formation plays an important role in wound colonization and infection, which impairs wound healing (6). The c-di-GMP monomer shows a two-fold symmetry, with two GMP moieties fused by a 5’-3’ macrocyclic ring (Figure 1.1 a). Cellular levels of c-di-GMP are regulated in response to internal and environment cues (3).
Recent studies revealed the importance of a novel hybrid cyclic dinucleotide, cyclic-GAMP. Bacterial c-GAMP exhibits 3’-3’ linkage and mammalian c-GAMP shows a 2’-3’ linkage (Figure 1.1 c & b respectively). 3’3’-GAMP is synthesized by the novel dinucleotide cyclase DncV in *Vibrio cholera*. This bacterial c-GAMP was shown to be responsible for efficient intestinal colonization and chemotaxis regulation in the bacterium *V. cholera* (7). 2’3’-cGAMP was found as a product of cGAMP synthase, cGAS, in mammals (8). This mammalian c-GAMP acts as second messenger to activate the innate immune responses (9, 10).

![Figure 1.1. Structure of cyclic dinucleotides. a) 3’5’-c-di-GMP b) 2’3’-c-GAMP c) 3’3’-c-GAMP](image)

The intracellular levels of CDNs are tuned by the concerted (and somewhat opposing) action of two types of enzymes: a) diguanylate synthases (CDN synthesis from two NTP molecules) and b) phosphodiesterases (CDN hydrolytic degradation to either the linear 5’phosphoguanylyl-(3’-5’)-guanosine (5’-pGpG) or GMP)). HD-GYP proteins represent a relatively novel addition to the PDE functional superfamily, and belong to a large superfamily of proteins annotated as HD-domain proteins. These HD-domain proteins contain the characteristic tandem histidine-aspartate structural motif. Here we are interested in the representative HD-GYP proteins that have two tandem HD binding motifs.
HD-GYP phosphodiesterases (PDEs)

HD-GYP consists of a large subfamily of proteins, which belongs to the larger HD-domain superfamily of metal-dependent hydrolases and oxygenases. HD-GYP proteins can come in a variety of types and flavors with respect to the domains they are composed of. Some HD-GYP proteins contain binuclear metal-binding site, while some can bind three metal cofactors at the active site. Some HD-GYP proteins contain an extra domain that was predicted as an activator of the PDE activity. HD-GYP proteins break down c-di-GMP to GMP via the linearized product 5’-pGpG, even though in some cases 5’-pGpG was reported as the only detected product (Figure 1.2) (11, 12)

![Diagram of c-di-GMP hydrolysis by HD-GYP phosphodiesterases.](image)

Figure 1.2. Scheme of c-di-GMP hydrolysis by HD-GYP phosphodiesterases.

There are only four X-ray crystal structures presently available of HD-GYP PDEs (Figure 1.3). All structures demonstrate that for the ligation of a dimetal site (at least) seven protein residues are needed that are arranged in the following motif: [H…HD…H…HH…D].

3
Figure 1.3. Crystal structures of HD-GYP protein active sites. a) crystal structure containing diiron binding sites 3.41 Å away in *Bdellovibrio bacteriovorus* Bd1817 (PDB entry 3TM8). b) crystal structure containing di-nickel binding sites 5.46 Å away in *Pseudomonas aeruginosa* PA4781 (PDB entry 4R8Z). c) crystal structure containing triiron binding sites 3.70 Å and 3.42 Å away respectively in *Persephonella marina* EX-H1 *PmGH* (PDB entry 4MCW). d) crystal structure containing diiron binding site 3.67 Å away in *Listeria monocytogenes* PgpH (PDB entry 4S1C). e) crystal structure of PgpH in complex with c-di-AMP (PDB entry 4S1B)

The first crystal structure is from the HD-GYP homolog from *Bdellovibrio bacteriovorus* (Bd1817), which however is inactive (has lost the ability to hydrolyze c-di-GMP) (13). The structure of Bd1817 reveals a diiron binding site of 3.41 Å away. The second structure of an HD-GYP protein was that of the PDE from *Pseudomonas aeruginosa*, namely PA4781. The sequence of PA4781 encodes upstream of the HD-GYP domain a REC domain (of a two-component regulator) (14). PA4781 exhibits very low activity and is purported to catalyze c-di-GMP in a two-step hydrolytic reaction. However, the crystal structure of the latter demonstrates di-metal site of Ni 5.46 Å away. Intriguingly, the third available structure of an HD-GYP protein involved in c-di-GMP hydrolysis showed an unusual triiron site, which was postulated to be coordinated by the presence of an additional glutamate ligand, which can be considered as a hallmark for the coordination of a trinuclear center. This triiron HD-GYP protein *PmGH* from *Persephonella*
*marina* EX-H1 showed complete hydrolysis of the c-di-GMP to GMP. The last crystal structure of an HD-GYP protein is in complex with substrate c-di-AMP. This HD-GYP protein PgpH from *Listeria monocytogenes* hydrolyzes c-di-AMP into 5’-pApA (33).

**The Vibrio cholera VCA0681 and Shewanella oneidensis SO3491**

*Vibrio cholera* is an aquatic organism that can transiently colonize human small intestine, which can cause fatal diarrhea to the host. In marine habitats, *V. cholera* is often found attached to other organisms in microbial communities called biofilms (15). Understanding the mechanisms used by *V. cholera* to control biofilm formation has been an ongoing research. Previous work with infant mouse model has shown that c-di-GMP inhibits the ability of *V. cholera* to effectively colonize the intestine (16).

*V. cholera* encodes nine HD-GYP proteins that were predicted to be enzymatically active and four of them were found to be active *in vivo* (17). VCA0681 has two domains containing two concerted HD-domain dimetal sites coordinated by seven ligand residues, however only the C-terminal domain is followed by the GYP sequence triad (considered important for CDN hydrolysis) (Figure 1.4). Therefore, the VCA0681 is better described as containing a double HD-[HD-GYP] bimetallic active site. It is considered the most active HD-GYP phosphodiesterase in *V. cholera* to degrade c-di-GMP into the linearized product 5’-pGpG (18). Expression of VCA0681 significantly increased motility of *V. cholerae* through 0.3% agar medium and decreased biofilm formation in rich medium on glass (19). It is also the first reported phosphodiesterase that can break down the novel hybrid cyclic dinucleotide, c-GAMP (2).
VCA0681 is very challenging to study biochemically, because it is hardly soluble unless expressed as a fusion with the maltose-binding protein (MBP) (20). In addition, after proteolytic cleavage of the solubility tag (MBP), the half-life stability of the protein is of the order of an hour, precluding any crystallization attempts. We generated a sequence similarity network and phylogenetic analysis to get a better insight into the HD-GYP superfamily proteins. VCA0681’s homolog, SO3491, is from the model γ-proteobacterium Shewanella oneidensis. We selected this protein on the basis of the sequence similarity network. SO3491 is 60% sequence similarity and 43% identity with VCA0681. It was also highly predicted for crystallization. Based on the bioinformatics analysis, SO3491 was chosen as a candidate that would be better suited for further biochemical and structural characterization.

Figure 1.4. A) A homology model of VCA0681 featuring both HD dinuclear centers. The color of each center correlates with the colored text. The GYP-triad is colored in green. B) Schematic construct of VCA0681 regarding two HD domains. C) Sequence alignment of the two individual HD domains in VCA0681, where HD1 represents HD75 (residues 1 – 233) and HD2 represents HD289 (residues 234 – 431). The aligned six residues at the binding sites are boxed in green. The sequence similarity and identity between these two HD domains are 18%, respectively.
Goal of the thesis

In my thesis project, the main goal is to understand the chemical nature of the transition metal content and the two dimetal domains in VCA0681 and its homolog SO3491, and dictate substrate specificity and reaction outcome (one- vs two-step hydrolysis). As described, we focused on the VCA0681 homolog SO3491, because it was better suited for us to characterize it both biochemically and structurally. To better understand the roles of individual HD domain, knockouts of each HD domain mutants were constructed and assessed for activities. Extensive bioinformatics and phylogenetic analysis were carried out, in which we collected HD-GYP proteins with scant biochemical information. The overall aim is to be able to gain insight into the chemical diversity and evolutionary diversification of the HD-GYP protein superfamily, with respect to the chemical nature of the cofactor, their respective specific activities and their reaction outcome. The invaluable information gained can serve as a functional map predictor that will allow portraying the differences in these proteins and allowing for concluding for their individual roles in the cell.
Chapter 2

Experimental Procedures

2.1 Material and methods

All chemicals unless specified were obtained by Fisher Scientific (NH) and were of high purity grade. c-di-GMP, 5’-pGpG and c-GAMP were initially purchased from Axxora, LLC (Farmingdale, NY), and GMP was from Acros Organics. In all subsequent experiments, the cyclic dinucleotides were enzymatically synthesized employing the WspR* and DncV cyclases, as previously described (21).

2.2 Plasmid construction, cloning and mutagenesis

pSUMO-SO3491

The gene encoding for the HD-GYP phosphodiesterase protein from *Shewanella oneidensis* was inserted into the pSUMO expression vector (LifeSensors, Inc) with kanamycin resistance to express the protein of interest as a linear chimera with the SUMO protein tag. The vector was
kindly gifted by Dr. Squire J. Booker (The Pennsylvania State University). This vector encodes for an N-terminal His6-tag, the small ubiquitin-like modifier protein (SUMO), and a Ulp1 recognition site for the subsequent cleavage of the SUMO-tag from the protein of interest.

**MBP-SO3491**

The SO3491 encoding gene was inserted with PCR-based cloning using the primers listed in Table 1 via the HindIII and XhoI restriction sites in the PDB.His.MBP (Berkeley Structural Genomics Center) expression vector that has kanamycin resistance. This vector encodes for an N-terminal His6-tag, the maltose binding protein (MBP), and TEV recognition site prior to the SO3491 gene, to allow for the subsequent cleavage of the MBP-tag.

**D69A and D283A pSUMO-SO3491**

Single-point amino acid substitutions were generated by back-to-back PCR (Q-5, New England Biolabs, MA) using the primers listed in Table 1.

**D69A MBP-SO3491**

The D69A SO3491 encoding gene was inserted with PCR-based cloning using the primers listed in Table 1 via the HindIII and XhoI restriction sites in the PDB.His.MBP (Berkeley Structural Genomics Center) expression vector that has kanamycin resistance.

**WT, D75A and D28A variants of VCA0681**

The plasmids expressing for the wild-type and single-point variants of VCA0681 as fusion proteins with an MBP-tag were kindly gifted by Dr. Donald M. Kurtz (Department of Chemistry, University of Texas at San Antonio) (9).
2.3 Protein expression and purification

All plasmids were transformed in T7 *Escherichia (E.) coli* express competent cells (New England Biolabs, MA). Transformed cells were grown in M9 minimal media (M9) to direct specific metal incorporation with either 200 µM Mn²⁺ or 250 µM Fe²⁺ with 50 mg/L kanamycin at 37 °C with shaking (220 rpm) until OD₆₀₀ reached a value between 0.6-0.8. Protein expression was then induced with the addition of 0.25 mM IPTG. Cell cultures were incubated at 18 °C with shaking (220 rpm) for 16-20 hours. Cultures were centrifuged at 8,000 x g for 20 minutes; cell pellets were flash frozen in liquid N₂ and stored at -80 °C. Cell pellets were resuspended in the lysis buffer (50 mM HEPES, 300 mM NaCl and 10 mM imidazole, pH 7.5). The suspension was lysed using a QSonica sonicator (Newton, CT) (~ 30 min per 100 mL suspension) with the addition of 250 µM ammonium iron sulfate and 0.045 mg/mL PMSF. Another 0.045 mg/mL PMSF was added after sonication. Lysed cells were centrifuged at 22,000 x g for 30 minutes. The supernatant was loaded onto a Ni²⁺-NTA immobilized affinity chromatography column (~100 mL resin per 500 mL lysate) equilibrated with the lysis buffer. The column was washed first with lysis buffer and then with buffer A (50 mM HEPES, 20 mM imidazole and 200 mM NaCl, pH 7.5). The protein was eluted by washing with buffer B (50 mM HEPES, 150 mM NaCl and 300 mM imidazole, pH 7.5). Fractions containing the protein were pooled and concentrated at 3,500 x g using a 30 K Amicon Ultra-15 Centrifugal Filter Unit (Tullagreen, Carrigtwohill Co. Cork, Ireland). The concentrated protein was then re-buffered in the storage buffer (50 mM HEPES, 300 mM NaCl and 10%, pH 7.5) using the same centrifugal device to remove imidazole. Protein for crystallography was further purified by size exclusion chromatography using a HiLoad¹⁴ 16/600 Superdex™ 200 pg column equilibrated with the storage buffer. Fractions containing the purified protein were pooled and concentrated at 3,500 x g using a 30 K Amicon Centrifugal Filter. The protein was flash frozen in
liquid $N_2$ and stored at -80 °C. Protein purity was estimated by SDS-PAGE with Coomassie staining, and protein concentration was determined by using a molar absorption coefficient of 48,820 M$^{-1}$•cm$^{-1}$ at 280 nm (https://web.expasy.org/protparam/). The iron content was determined by the ferrozine assay.

2.4 Cleavage of the SUMO tag by the Ulp1 protease

The ubiquitin-like protease 1 (Ulp1), recognizes the three-dimensional structure of the SUMO domain and cleaves after di-glycine at the C-terminus. For the protease cleavage reactions, Ulp1 (~1 µg per mg of pSUMO-SO$_3491$) was incubated with SUMO-SO$_3491$ for 16 to 20 hours on ice. The reaction mixture was then loaded onto a Ni$^{2+}$-NTA immobilized affinity chromatography column (~100 mL resin per 500 mL lysate) equilibrated with the lysis buffer. The cleaved proteins were eluted by washing the column with the lysis buffer. The SUMO-tag, Ulp1 and uncleaved pSUMO-SO$_3491$ (which all contain an N-terminal His-tag) were then eluted with buffer B. Fractions containing the untagged (native) SO$_3491$ protein were pooled and concentrated at 3,500 x g using a 30K Amicon Centrifugal Filter. The concentrated protein was re-buffered in the storage buffer and concentrated to the desired final concentration. Protein for crystallography was further purified by size exclusion chromatography using a Sephadex S-200 column equilibrated with 50 mM HEPES Buffer (pH 7.5) containing 300 mM NaCl, 10 mM imidazole and 10% glycerol. Fractions containing the protein were pooled and concentrated at 3,500 x g using a 30 K Amicon Centrifugal Filter. The protein was flash frozen in liquid $N_2$ and stored at -80 °C. Protein purity was estimated by SDS-PAGE with Coomassie staining, and protein concentration was determined
by using a molar absorption coefficient of 45,540 M$^{-1}$cm$^{-1}$ at 280 nm. The iron content was determined by the ferrozine assay.

2.5 Cleavage of the MBP-tag with the S219V TEV protease

The tobacco etch virus (TEV) protease (27 kDa) recognizes the ENLYFGQ sequence and cleaves between Q and G. The cleavage reactions were initiated by addition of TEV (~ 1 µg per mg of MBP-SO$_{3491}$) and they were carried out for 4 hours on ice. A second addition of the same amount of TEV is performed and the reactions if further incubated for another 12 to 16 hours on ice (total ratio of TEV to MBP-SO$_{3491}$ was 1:25. The reaction mixture was then loaded on to a MBP-trap HP prepacked column (GE Healthcare) equilibrated with the storage buffer. The overnight reaction mixture was loaded onto the column with a flow rate of 0.5 mL/min. Fractions containing the untagged (native) SO$_{3491}$ were collected by washing the column with the storage buffer and concentrated using a 30 K Amicon Centrifugal Filter. TEV and uncleaved MBP-SO$_{3491}$ (both of which contain an N-terminal His$_6$-tag) were eluted by washing the column with elution buffer containing 50 mM HEPES, 150 mM NaCl, 10mM Maltose and 10% glycerol (pH 7.5). Protein purity was estimated by SDS-PAGE with Coomassie staining. Due to the low binding affinity of the MBP-tag of the fusion protein, the first purification could not provide pure cleaved SO$_{3491}$. The collected fractions were further purified by size exclusion chromatography using a Sephadex S-200 column equilibrated with the storage buffer. Fractions containing the cleaved protein were pooled and concentrated at 3,500 x g using a 30K Amicon Centrifugal Filter. Protein purity was estimated by SDS-PAGE with Coomassie staining again. The concentrated protein mixture was then purified a second time with MBP-trap HP prepacked column (GE Healthcare). Protein purity was again estimated by SDS-PAGE with Coomassie staining.
2.6 Ferrozine assay

The Fe content was determined by ferrozine assay following the procedures of Bollinger (Bollinger, 1993).

The protein was diluted to 150 µM with water to a final volume of 83.3 µL. After the addition of 16.7 µL of 50% TCA, the sample was spun at 17,000 x g for 2 min. The volume of the supernatant was carefully measured and the supernatant was mixed with 20 µL 75 mM ascorbic acid, 20 µL 10 mM ferrozine, 120 µL saturated NH₄OAc and water to make a 660 µL solution. To make a control sample, 500µL water was used without any protein. The Fe concentration was estimated on the basis of the absorbance at 562 nm, corresponding to the Fe²⁺-ferrozine chromophore.

2.7 High Performace liquid chromatography (HPLC)

Denatured reaction samples were centrifuged at 21,130 x g for 2 minutes and the supernatant were filtered with 0.22 µm nylon Spin-X Centrifuge Tube Filter (Corning Incorporated, Corning, NY). The filtered samples were then transferred into HPLC tubes analyzed on a 6125B Agilent spectrometer equipped with a 1260 Infinity Liquid Chromatography system and a 1260 Infinity Photodiode Array Detector WR. Samples were injected on an Agilent reverse-phase C18-A Polaris column (particle size 5 µm, 150 x 4.6 mm) and educts and products were separated using a gradient method utilizing a water-based mobile phase (10 mM KH₂PO₄ and 10 mM TBAH, pH 6, Solvent A) and an organic-based mobile phase (methanol with 10 mM TBAH, Solvent B). Substrates and products were eluted from the column with a gradient of 95% solvent A and 5% solvent B to 50%
solvent A and 50% solvent B at a flow rate of 1.5 mL/min for 25 min. Comparison of integrated peak intensities to that of internal standards (substrates and products) of known concentration enabled quantification of the analyses. Nucleotides were detected at a wavelength of 254 nm.

2.8 Multiple turnover assays

End-point reactions to test for activity of the wild-type and variant HD-GYP domain proteins with substrates and substrate analogues contained final concentrations of 1 μM (di-iron concentration) wildtype and knockout-construct SO_3491 and VCA0681 for c-di-GMP reactions and 5 μM (di-iron concentration) for c-GAMP reactions, 100 μM MgCl₂, 200 μM sodium dithionite and 50 μM substrate in 100 mM HEPES buffer (pH 8) containing 150 mM NaCl and 10% glycerol. Reactions examining the substrate specificity of the HD-GYP proteins were carried out with 5 μM (di-iron concentration) SO_3491, 100 μM MgCl₂, 200 μM sodium dithionite and 50 μM of substrates (c-diAMP, 5′pGpG, 5′pApG and 3′5′cGMP) in 100 mM HEPES buffer (pH 8) containing 150 mM NaCl and 10% glycerol. All reaction components (O₂-free) were mixed in an anoxic chamber (CoyLab). Reaction buffer, MgCl₂, enzyme and sodium dithionite (50-fold excess with respect to cofactor concentration) were incubated for 20 minutes before the addition of substrate to allow full reduction of the metal cofactor to the diferrous form (as confirmed in ⁵⁷Fe-labeled samples examined by Moessbauer spectroscopy, which confirm that the cofactor is in the Fe₂⁺/II form). All reactions were initiated with the addition of substrate or substrate analogs and terminated at the described time-points by quenching at 95 °C for 10 min. The efficiency of heat quenching for rapid inactivating the enzymes at the desired time-points, were confirmed by comparison of quenching of the reactions in 2 M acetic acid.
2.9 Determination of steady-state parameters of SO3491 with c-di-GMP and c-GAMP

Reactions to test for activity of the native SO_3491 with c-di-GMP contained final concentration of 0.4 µM (Fe$^{2+}$ concentration) wild-type untagged SO_3491 and varying concentrations of c-di-GMP (5 µM, 10 µM, 20 µM, 60 µM, 120 µM and 240 µM). The reaction mixture additionally contained 100 µM MgCl$_2$, 200 µM sodium dithionite and 50 µM substrate in 100 mM HEPES buffer (pH 8) containing 150 mM NaCl and 10% glycerol. All reaction components (O$_2$-free) were mixed in an anoxic chamber. Reactions were terminated at different time points by denaturing at 95 °C for 10 min. Initial rates of reactions were obtained by fitting concentrations of c-di-GMP over time with a linear equation. The initial rates for different c-di-GMP concentrations were then plotted and fitted with a Michaelis-Menten equation to determine $k_{cat}$ and $K_M$ according to the equation:

$$v = \frac{v_{max} \cdot [c-di-GMP]}{K_M + [c-di-GMP]}$$  \hspace{1cm} (Eq. 1)

2.10 UV-VIS spectroscopy

UV-visible spectra were recorded on a Cary 60 spectrometer (Varian, Walnut Creek, CA) using the associated WinUV software package.
2.11 Thermofluor assays

The apparent melting temperature values for the HD-GYP proteins in this study were determined by performing fluorescence thermal shift assays on a StepOne Plus Real Time PCR System (Company, city). Experiments were carried out in MicroAmp Fast Optical 96-Well plates, in 25 µL reactions including 10x SYPRO fluorescent dye, 10 µM protein in a buffer containing 50 mM HEPES, pH 8.0, and 200 mM NaCl. Reactions were heated from 25 to 95 °C, and fluorescence intensity (570 nm) was recorded at 0.3 °C temperature increments.

2.12 Mössbauer Spectroscopy

All samples were prepared in storage buffer under oxygen-free conditions in an anaerobic glovebox (Coylab). The samples were reacted with an excess amount of sodium dithionite (10-20 equivalents) for 20-30 minutes at 22 °C prior to freezing in liquid N₂. Mössbauer spectra were recorded on WEB Research (Edina, MN) instruments that have been described previously (15). The spectrometer used to acquire the weak-field spectra is equipped with a Janis SVT-400 variable-temperature cryostat. The external magnetic field was applied parallel to the γ beam. All isomer shifts are quoted relative to the centroid of the spectrum of α-iron metal at room temperature. Mössbauer spectra were simulated using the WMOSS spectral analysis software (www.wmoss.org, WEB Research, Edina, MN).
2.13 Electron paramagnetic resonance (EPR) spectroscopy

All samples were prepared in storage buffer under oxygen-free conditions in an anaerobic glovebox (Coy). The samples were reacted with ascorbate sodium salt for 30 minutes at 22 °C prior freezing in liquid N₂. EPR spectra were acquired on a Bruker E500 Elexsys continuous wave (CW) X-Band spectrometer (operating at approx. 9.38 GHz) equipped with a rectangular resonator (TE102) and a continuous-flow cryostat (Oxford 910) with a temperature controller (Oxford ITC 503). The spectra were recorded at variable temperatures between 10-40 K at a microwave power of 0.2 mW, using a modulation amplitude of 1 mT, a microwave frequency of 9.38 GHz.
**Table 1.** List of plasmid and primers

<table>
<thead>
<tr>
<th>Protein</th>
<th>Vector</th>
<th>Restriction Sites</th>
<th><strong>Forward Primer (5’ – 3’)</strong></th>
<th><strong>Reverse Primer (3’ – 5’)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>PDB.His.MBP</td>
<td>HindIII/XhoI</td>
<td>AATA AAG CTT ATG AAC GCG TTC CTG</td>
<td>ACAA CTC GAG TCA GGT TTG AAA</td>
</tr>
<tr>
<td>D69A</td>
<td>pSUMO</td>
<td>EcoRI/HindIII</td>
<td>CTG ATC CAC GCG TGC GGT GTG</td>
<td>ACC CGC AAA ATA CGC GAA CTG</td>
</tr>
<tr>
<td>D283A</td>
<td>pSUMO</td>
<td>EcoRI/HindIII</td>
<td>CTG CTG CAC GCG GTT GGC AAG</td>
<td>ACC CGC CAC GTC CAG CAG</td>
</tr>
</tbody>
</table>

**Table 2.** Protein samples and metal content

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Vector</th>
<th>Metal Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt SO3491</td>
<td>pSUMO</td>
<td>1.8 Fe/Protein</td>
</tr>
<tr>
<td>wt SO3491</td>
<td>MBP</td>
<td>1.6 Fe/Protein</td>
</tr>
<tr>
<td>D69A SO3491</td>
<td>MBP</td>
<td>0.96 Fe/Protein</td>
</tr>
<tr>
<td>D283A SO3491</td>
<td>pSUMO</td>
<td>1.2 Fe/Protein</td>
</tr>
<tr>
<td>wt VCA0681</td>
<td>MBP</td>
<td>2.6 Fe/Protein</td>
</tr>
<tr>
<td>D75A VCA0681</td>
<td>MBP</td>
<td>1.1 Fe/Protein</td>
</tr>
<tr>
<td>D289A VCA0681</td>
<td>MBP</td>
<td>1.3 Fe/Protein</td>
</tr>
</tbody>
</table>
Chapter 3

Results

The non-heme carboxylate-bridged diiron cofactor in HD-domain proteins can attain three overall redox states, differing in the level of reduction, namely Fe\textsuperscript{III}Fe\textsuperscript{III} (diferric), Fe\textsuperscript{II}Fe\textsuperscript{III} (mixed-valent) and Fe\textsuperscript{II}Fe\textsuperscript{II} (diferrous). These three redox states have been demonstrated for HD-domain oxygenases (i.e. myo-inositol oxygenase- MIOX, and the organophosphonate degrading PhnZ), for which the mixed-valent (MV) form of the cofactor can be accumulated to ~60-70%. This is in contrast to canonical diiron oxygenases and oxidases of the ferritin-like superfamily, which do not detectably accumulate the MV form. This discrepancy can be rationalized by the fact that ferritin-like non-heme dinuclear sites activate their substrates using O\textsubscript{2} from the diferrous manifold, whereas HD-domain oxygenases catalyze their four-electron oxidative conversions from the MV manifold. Diiron hydrolases are not that common, but it has been proposed that the hydrolysis reaction occurs at the Fe\textsuperscript{II}Fe\textsuperscript{II} level of the cofactor. It is presently not well understood whether HD-domain hydrolytic enzymes can accumulate the Fe\textsuperscript{II}Fe\textsuperscript{III} form to significant yields and whether the MV form is potent towards hydrolysis.

VCA0681 has an unconventional arrangement, harboring two diiron sites, and has been shown to be strictly Fe-dependent (no other metal ions could stimulate activity, with the exception
perhaps of Mn$^{2+}$, which allowed a detectable but otherwise very sluggish activity to be considered relevant/physiological) (12). Only the C-terminal diiron site is an active PDE domain (on the basis of knock-out variants of VCA0681) (12), whereas the role/function of the other bimetallic site is presently unknown.

3.1 Wildtype VCA0681 and SO3491

We proceeded to study the active-sites of wt MBP-VCA0681 and wt MBP-SO3491 by Mössbauer and EPR spectroscopies, so as to 1) characterize the individual forms of the enzymes we would carry out activity assays with; 2) to ensure that the Fe in the sample is assembled exclusively in bimetallic sites and does not exists in monomeric species (which would not support activity); 3) and lastly quantify the extent of accumulation of the Fe$^{II}$Fe$^{III}$ and compare it to that stabilized in HD-domain oxygenases (i.e. PhnZ) (22).
Figure 3.1 Mössbauer spectra of wt MBP-VCA0681 (left) and wt MBP-SO3491 (right) at three different redox states and in the absence of an external applied magnetic field. The Mössbauer spectra of wt MBP-VCA0681 were measured at 120 K. The Mössbauer spectra of wt MBP-SO3491 were measured at 80 K. At both temperatures, electronic relaxation is much faster than the nuclear larmor frequency, and as a result both paramagnetic states and diamagnetic states produce quadrupole doublets.

The wt MBP-VCA0681 and wt MBP-SO3491, demonstrated somewhat different Mössbauer spectra in their as-purified state (Figure 3.1). The spectrum of the wt VCA0681 can be described by two quadrupole doublets, with parameters characteristic of anti-ferromagnetically (AF)-coupled Fe\textsuperscript{III} sites in diiron centers ($S = 0$). The two doublets with average Mössbauer parameters: ($\delta_1 = 0.48$ mm/s, $\Delta E_{Q1} = 0.88$ mm/s and $\delta_2 = 0.54$ mm/s, $\Delta E_{Q2} = 1.80$ mm/s) are not in 1:1 ratio. This is particularly more pronounced in the spectrum of the MBP-SO3491, suggesting that this site differentiation could either correspond to the different HD-domains in the protein (N-terminal and C-terminal), or to a mixed population of active site molecules with a slight deviant
ligation or the presence of a small-ligand bound to the active site. The two last hypotheses, are perhaps more plausible when compared to the Mössbauer data of another, as-purified HD-GYP that harbors only one diiron site (e.g. VCA0931, results not shown). Different levels of reduction were achieved by reducing either with sodium ascorbate (weak reducing agent) or sodium dithionite (strong reducing agent). Addition of excess of sodium ascorbate (5 mM) under O₂-free conditions produced also different extent of reduction both for VCA0681 and SO3491 (Figure 3.1). Quantification of the respective redox states yields for the VCA0681: 10% Fe³⁺Fe³⁺ (all-ferric), 45% Fe²⁺Fe³⁺ (MV), and 45% of Fe²⁺Fe²⁺ (all-ferrous), while for the SO3491 the distribution is somewhat different with: 65% Fe³⁺Fe³⁺ (all-ferric), 16% Fe²⁺Fe³⁺ (MV), and 19% of Fe²⁺Fe²⁺.(all-ferrous), respectively. This is in contrast to oxygenases, for which under these mild reducing conditions almost no diferrous form accumulates (<4%) and the majority of the sample is in the MV form (S = 1/2, ~70%) and the rest in a diferric form (S = 0, ~30%) (23). This result demonstrates that the two redox couples, are now ‘closer’ to each other and that there is an upshift in the reduction potential of the Fe²⁺Fe³⁺ /Fe²⁺Fe²⁺ couple as demonstrated by its significant accumulation under mild reducing conditions.
Figure 3.2. Mössbauer spectra of the MBP-SO3491 reduced with 5 mM sodium ascorbate (left) and 10 mM sodium dithionite (right). Doublets colored in varying colors correspond to the relative amount of the iron in each redox state. **Left:** sodium ascorbate reduced MBP-SO3491 gave 16% MV form, 19% diferrous form and the rest in the diferric form. The green shade corresponds the FeII amount complexed in the Fe^II-Fe^III state, the blue doublet to the Fe complexed in the Fe^II-Fe^II state, the red shade to the Fe^III-Fe^III form, and the orange shade to the Fe^II and is part of the Fe^II-Fe^III state. **Right:** sodium dithionite reduced MBP-SO3491 gave 100% diferrous form. The spectrum was fitted with two quadrupole doublets corresponding to two Fe^II-Fe^III species. The spectra were recorded at 80 K and in the absence of an applied magnetic field.

These two combined observations are in contrast to those observed for HD-domain oxygenases, for which the yield of the Fe^II-Fe^III form is the same irrespective of the excess sodium ascorbate level. It has to be noted, that we performed reduction with sodium ascorbate under the exact same conditions with all three proteins (MBP-VCA0681, MBP-SO3491, and the PhnZ oxygenase). The rationale for using a limiting amount of reductant was to be able to accumulate Fe^II-Fe^III and not as much (if at all) Fe^II-Fe^II, so that we could carry out activity assays to demonstrate whether the mixed-valent form is active. On the other hand, reduction of both wt MBP-VCA0681 and MBP-SO3491 with excess of sodium dithionite under O₂-free conditions, leads to 100% reduction of the diiron sites to the diferrous state (Figure 3.2).
Figure 3.3. CW EPR spectra of the different forms of SO3491 and MBP-VCA0681. All samples were reduced with 5 mM sodium ascorbate and recorded at 10 K. Experimental conditions: microwave frequency = microwave power = 0.2 mW, modulation amplitude: 1mT.

The estimation of the extent of accumulation of the MV form was additionally supported by recording EPR spectra on parallel Mössbauer samples, and quantifying the extent of accumulation of Fe$^{II}$Fe$^{III}$ (it is the only EPR-active state ($S = 1/2$), whereas the other two forms are EPR-silent). The continuous wave (CW) EPR spectra of MBP-VCA0681 and SO3491 demonstrated the characteristic signals of anti-ferromagnetically (AF)-coupled Fe$^{II}$Fe$^{III}$ ($S = 1/2$) centers, with $g_{av} < 2$ (Figure 3.3A). The three principal g-values of the wt MBP-VCA0681
spectrum are 1.94, 1.82 and 1.7 and those of MBP-SO3491 are 1.93, 1.8, 1.65, respectively. The slight shifts in the g-values and the broadness of the spectra are representative of slight differences in the local active-site protein environment and/or structural heterogeneity of their active centers. The wt SUMO-SO3491, MBP-SO3491 and native (“untagged”) SO3491 all demonstrate the same features in their EPR spectra (Figure 3.3B). We conclude that the tag does not alter the properties or the local environment surrounding the two bimetallic sites. The intensity of SUMO-SO3491 were significantly smaller than that of MBP-VCA0681, which is consistent with the Mössbauer quantification. MBP-VCA0681 accumulated more $\text{Fe}^{II}\text{Fe}^{III}$ than MBP-SO349 when reduced under same conditions (Figure 3.3C).

**Redox State Differentiation**

After establishing the distribution of the respective diiron redox states in the samples, we proceeded to examine whether the MV form is also active towards c-di-GMP hydrolysis. We thus performed activity assays under conditions of the ‘same’ cofactor concentration, on the basis of both the EPR and Mössbauer quantifications. Our activity assays were performed with sodium ascorbate reduced and sodium dithionite reduced proteins and with c-di-GMP as a substrate in a time-dependent manner. Both protein samples showed accumulation of the linearized product 5’-pGpG over time (Figure 3.4A, B). The peak areas of c-di-GMP and 5’-pGpG were obtain after integration of the HPLC elution profile (Figure 3.4C, D). On the basis of the EPR and Mössbauer quantifications, there is ~1 µM $\text{Fe}^{II}\text{Fe}^{III}$ and ~1.2 µM $\text{Fe}^{II}\text{Fe}^{II}$ present in the sodium ascorbate reduced sample. If both forms (MV and diferrous) of MBP-SO3491 are active, the effect on product formation would be additive. However, the yield of product (5’-pGpG) by the ascorbate reduced MBP-SO3491 is comparable to that obtained in the case of the dithionite reduced enzyme.
that contains 1 μM differrous. Thus, we conclude that the mix-valent MBP-SO3491 is nearly not as active towards hydrolysis (if at all). To unequivocally substantiate our results, we will perform single-turnover assays, which would provide an additional direct way to allow for assessing the catalytic competence of the MV form. Therefore, all assays reported hereafter, were carried out with dithionite-reduced samples that are 100% enriched in the Fe$^{II}$Fe$^{II}$ form, facilitating the estimation of the active cofactor in our assays.

Figure 3.4. Time-dependent activity assays with ascorbate and dithionite reduced MBP-SO3491 using c-di-GMP as a substrate. A, C) wild-type MBP-SO3491 was reduced with same cofactor to
ascorbate ratio for as the EPR sample of 20 minutes under O₂-free conditions. B, D) wild-type MBP-SO3491 was reduced with 200 equivalent cofactor concentrations of NaDT for 20 minutes under anaerobic condition. The reactions were heat-quenched at 15 seconds, 30 seconds and 1 minute inside the glovebox.

Activities of the wt MBP-VCA0681 and wt MBP-SO3491 were assayed and compared for the two different substrates c-di-GMP and c-GAMP (Figure 3.5). Whereas both CDNs have been shown previously as substrates of VCA0681 (both in vivo and in vitro), there was no prior evidence that the SO3491 encodes for an enzymatically active PDE or that it can process both substrates. Intriguingly, wt MBP-SO3491 can degrade c-di-GMP faster than the wt MBP-VCA0681 under the same conditions (experiments were performed in duplicates). In contrast, wt MBP-VCA0681 is shown to break down c-GAMP to 5’-pApG faster than the wt MBP-SO3491 (Figure 3.5).
Figure 3.5. Activity assays with NaDT reduced wt MBP-SO3491 and wt MBP-VCA0681 using c-di-GMP and c-GAMP as substrates. A, C) 1 µM (Fe²⁺) wt MBP-SO3491 and wt MBP-VCA0681 were reacted with 50 µM c-di-GMP for 4 minutes, 8 minutes, 1 hour and 2 hours. Dark green bars represent c-di-GMP and light green bars represent 5’-pGpG. B, D) 5 µM (Fe²⁺) wt MBP-SO3491 and wt MBP-VCA0681 were reacted with 50 µM c-GAMP for 10 minutes, 1 hour and 2 hours, respectively. Orange bars represent c-GAMP and yellow bars represent 5’-pApG. All assays were carried out at room temperature under O₂-free conditions in an anaerobic glovebox.

**Metal Cofactor Specificity**

Manganese was another potential metal ion that could yield detectable yet very slow PDE activity (12). Here we expressed and purified MBP-SO3491 with controlled manganese
supplementation in order to examine the PDE activity of the Mn-loaded MBP-SO3491. The manganese content of Mn-MBP-SO3491 could not yet precisely been determined (the ICP-AES results are pending). All experiments with the Mn-MBP-SO3491 activities were carried out assuming 2 Mn/protein on the basis of the analogous Fe content with Fe-MBP-SO3491. To avoid any contaminating activity due to the residual Fe co-purified with the Mn-MBP-SO3491 proteins (0.45 Fe/protein), we performed the assays under aerobic conditions, with samples that were prior incubated with 1 equivalent amount (with respect to cofactor concentration) of FeKCN₆ for 20 minutes before addition of the substrate. FeKCN₆ will fully oxidize any residual reduced forms of the cofactor to the diferric form, which is inactive. Figure 3.6 shows the activities of MBP-SO3491 with different metal cofactors and their abilities to use c-di-GMP and c-GAMP as substrates in a time-dependent manner. Although a lot slower than the Fe-loaded MBP-SO3491, Mn-loaded MBP-SO3491 can also degrade c-di-GMP and c-GAMP into 5’-pGpG and 5’-pApG.

Figure 3.6. Activity assays with Mn-loaded wild-type MBP-SO3491 with c-di-GMP and c-GAMP as substrates. A) 1 μM (dimanganese) wt MBP-SO3491 were reacted with 50 μM c-di-GMP for 4 minutes, 8 minutes, 1 hour and 2 hours. Dark green bars represent c-di-GMP and light green bars represent the 5’-pGpG product. B) 5 μM (dimanganese) wt MBP-SO3491 were reacted with 50 μM c-GAMP for 10 minutes, 1 hour and 2 hours. Orange bars represent c-GAMP and yellow bars
represent the 5’-pApG product. These assays were performed at room temperature, under air and in the presence of FeKCN.

Substrate Specificity

c-di-GMP and c-GAMP are reported as physiological substrates for VCA0681. MBP-SO3491 was assayed using seven nucleotides to probe its substrate specificity. C-di-GMP and c-GAMP were used as controls. 5’-pGpG and 5’-pApG, the linearized dinucleotides, were assayed to examine whether wild-type MBP-SO3491 can carry out the second step hydrolysis. C-di-AMP and 3’5’-cGMP were reported as substrates for other PDEs and here we assessed them as potential substrates for MBP-SO3491. We found that the PDE activity of MBP-SO3491 is limited to c-di-GMP and c-GAMP (Figure 3.7A). Wild-type MBP-SO3491 showed no activity with 5’-pGpG and 5’-pApG, which confirms its inability to carry out the second-step hydrolysis (to two molecules of GMP). Some activities were also observed with 2’3’-cAMP for SUMO-SO3419 (Figure 3.7).
Figure 3.7. A) Activity assays of wt MBP-SO3491 with different substrates (the c-di-AMP result not shown here). All activities were performed with 5 µM metal cofactor and 50 µM substrate analogs. The reactions were heat quenched after 1 hour. B) Activity assay of the wt SUMO-SO3491 (5 µM metal cofactor) with 50 µM 2’3’-cAMP. The reaction was heat-quenched after 1 hour.
Steady-State kinetic parameters for c-di-GMP and c-GAMP

Wild-type native (tagless) SO3491 was assayed with varying concentrations of c-di-GMP and c-GAMP, so as to examine the catalytic efficiency of this type of HD-GYP for the first time. The $k_{\text{cat}}$ and $K_M$ for c-di-GMP were determined to be of $8.7 \, \text{min}^{-1}$ and $3.1 \, \mu\text{M}$ respectively (Figure 3.5 A). The catalytic proficiency, $k_{\text{cat}}/K_M$, was calculated to be $4.7 \times 10^3 \, \text{M}^{-1} \, \text{s}^{-1}$ (Table 2). The cellular levels of c-di-GMP in *V. cholera* are reported to be within the range of 0.9 to 20 µM (24), which is consistent with our $K_M$ value. The $k_{\text{cat}}$ and $K_M$ for c-GAMP were determined to be of $0.32 \, \text{min}^{-1}$ and $55 \, \mu\text{M}$ respectively (Figure 3.8 B).

Figure 3.8. Steady-state kinetic parameters of SO3491 with c-di-GMP and c-GAMP as substrates. A) Wild-type native (‘untagged’) SO3491 (0.4 µM diiron) reduced with NaDT was assayed with 5 µM, 10 µM, 20 µM, 40 µM, 60 µM 120 µM and 240 µM c-di-GMP. B) wild-type native SO3491 (5 µM diiron) reduced with NaDT were assayed with 12.5 µM, 25 µM, 50 µM, 150 µM and 300 µM c-GAMP, respectively. All assays were carried out at room temperature under O$_2$-free conditions in an anaerobic glovebox, and the reactions were heat quenched at 100 °C.
Table 3. Kinetic parameters of cyclic dinucleotide specific phosphodiesterases belonging to the HD-GYP and the EAL subtype.

<table>
<thead>
<tr>
<th>PDE</th>
<th>Catalytic domain</th>
<th>CDN</th>
<th>$K_M$ (µM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_M$ (M$^{-1}$s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO_3491</td>
<td>HD-GYP</td>
<td>c-di-GMP</td>
<td>3.1</td>
<td>0.15</td>
<td>$4.7 \times 10^4$</td>
</tr>
<tr>
<td>SO_3491</td>
<td>HD-GYP</td>
<td>c-GAMP</td>
<td>55</td>
<td>$5.3 \times 10^{-3}$</td>
<td>97</td>
</tr>
<tr>
<td>PA4108$^1$</td>
<td>HD-GYP</td>
<td>c-di-GMP</td>
<td>20 ± 5</td>
<td>$1.5 \pm 0.1 \times 10^{-4}$</td>
<td>7.5</td>
</tr>
<tr>
<td>PA4781$^1$</td>
<td>HD-GYP</td>
<td>c-di-GMP</td>
<td>119 ± 30</td>
<td>$2.0 \pm 0.3 \times 10^{-4}$</td>
<td>2.7</td>
</tr>
<tr>
<td>PdeB ($Borrelia$ burgdorferi)$^{22}$</td>
<td>HD-GYP</td>
<td>c-di-GMP</td>
<td>0.0029</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>CC3396 ($Caulobacter$ crescentus)$^{25}$</td>
<td>EAL</td>
<td>c-di-GMP</td>
<td>&gt;100</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>RocR ($Pseudomonas$ aeruginosa)$^{26}$</td>
<td>EAL</td>
<td>c-di-GMP</td>
<td>3.2</td>
<td>0.67</td>
<td>$2.1 \times 10^5$</td>
</tr>
<tr>
<td>PA2567 ($Pseudomonas$ aeruginosa)$^{27}$</td>
<td>EAL</td>
<td>c-di-GMP</td>
<td>5.2</td>
<td>0.39</td>
<td>$7.5 \times 10^4$</td>
</tr>
<tr>
<td>VieA ($Vibrio$ cholera)$^{28}$</td>
<td>EAL</td>
<td>c-di-GMP</td>
<td>0.06</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>BlrP1 ($Klebsiella$ pneumoniae)$^{29}$</td>
<td>EAL</td>
<td>c-di-GMP</td>
<td>n.d.</td>
<td>0.13</td>
<td>n.d.</td>
</tr>
</tbody>
</table>

$O_2$-sensitivity of the ‘active’ Fe$^{II}$Fe$^{II}$ form of the wt MBP-SO3491

Wild-type MBP-SO3491 was fully reduced to the diferrous form with excess NaDT under $O_2$-free conditions. Dithionite was then removed from the samples by desalting under $O_2$-free conditions in the anaerobic chamber. The reduced protein was then reacted against different concentrations of an $O_2$-containing buffer (Figure 3.9).
The diferrous form of the protein is optically transparent. After reacting with O$_2$, oxidation to the Fe$^{III}$Fe$^{III}$ form is signified by an increase in the absorbance at 330 nm and 360 nm, which correspond to characteristic ligand-to-Fe$^{III}$ charge transfer transitions in non-heme oxo/carboxylate-bridged diferric (Fe$^{III}$Fe$^{III}$) sites (30). The kinetics of the oxidation of SO3491, were followed by monitoring the absorbance at 360 nm over time and these were fit using a double-exponential function according to the equation:

$$\Delta A_{365}(t) = \Delta A_1 \left(1 - e^{-k_1 t}\right) + \Delta A_2 \left(1 - e^{-k_2 t}\right) \quad (\text{Eq. 2})$$

, which gives the change in absorbance with respect to time for two irreversible first-order reactions as a function of the apparent first-order rate constants $k_1$ and $k_2$ (Figure 3.10B). The plot of the fastest apparent first-order rate constant ($k_1$) versus [O$_2$] yields a bimolecular rate constant (slope) $k(O_2) = 20$ mM$^{-1}$s$^{-1}$ (Figure 3.10C).
Figure 3.10. A) Time-dependent UV-VIS traces monitoring the spectral changes upon reaction of 180 µM diferrous wt MBP-SO3491 with 600 µM O₂. The absorbance at 330 nm and 365 nm increases with increasing reaction time denoting oxidation to the Fe³⁺Fe³⁺ form. B) The kinetics of the oxidation of the Fe²⁺Fe²⁺ cofactor to the Fe³⁺Fe³⁺ upon reaction with varying concentrations of O₂ show a multi-phasic behavior. C) Plot of the fastest apparent first order rate constant for [O₂] of 300 µM, 600 µM and 1200 µM. The plot was fitted with linear regression that yields a second-order rate constant of 20 mM⁻¹s⁻¹.

3.2 Individual HD-domains

Knockout constructs of each HD-domain were expressed and purified. For the SUMO-SO3491, both variants (D283A and D69A) showed no PDE activity (data not shown here).
However, based on previous results on D75A MBP-VCA0681 (ours and ref. 12), we expected that the analogous variant D69A (SUMO)-SO3491 should retain PDE activity (or be as active as the wt). To confirm that the aminoacid substitution was not destabilizing the structure of the protein, we replaced the N-terminal SUMO tag (12 kDa) with an MBP tag (42 kDa, similar to VCA0681). The protein stability was examined by a thermoshift assay.

We observed a dramatic decrease in the stability of the D69A SUMO-SO3491 (Figure 3.11A) compared to that of the wt SUMO-SO3491, suggesting that this single aminoacid substitution, greatly destabilizes the protein (most likely due to loss of the cofactor in the N-terminal site). The $\Delta T_m$ between wt and D69A was 8 °C. We reverted by PCR the D69A-SO3491 variant back to the wt SUMO-SO3491(A69D) and we could regain both PDE activity and protein stability, suggesting that our original substitution did not cause any other errors or artifacts (Figure 3.11). This result clearly demonstrated that the decrease in stability was indeed caused by the D to A substitution, leading to the observed misleading inactivity of the D69A variant (when compared to the active D75A cognate variant in MBP-VCA0681). As a next step, we proceeded to replace the small SUMO tag, with the bulkier MBP, with the aim to increase the stability of the protein, in particular in the N-terminal domain. We introduced the MBP-tag in both the wt and the D69A variants, purified the proteins and carried out thermoshift denaturation assays (Figure 3.11). The presence of the MBP tag indeed greatly stabilized the D69A with respect to the wt construct, as reflected in their smaller $\Delta T_m = \sim 1.5$ °C. To validate our hypothesis that the apparent loss of PDE activity is linked to the decreased protein stability, we proceeded to carry out activity assays with the D69A MBP-SO3491 variant, which was now shown to be active (as anticipated, *vide infra*).
Figure 3.11. Thermoshift denaturation curves of the different SO3491 constructs. A) Fluorescence intensities of the wt SUMO-SO3491, D69A-to-A69D SUMO-SO3491 (wt-like) and D69A SUMO-SO3491. B) Fluorescence intensities of the wt SUMO-SO3491, D69A SUMO-SO3491, wt MBP-SO3491 and D69A MBP-SO3491 over a range of temperatures (30 - 95 °C).

Knock-out constructs of the second HD-GYP domain for D289A MBP-VCA0681 and D283A MBP-SO3491 exhibited no PDE activity. Knock out construct of the first HD-domain D75A MBP-VCA0681 completely broke down c-di-GMP into GMP in 20 hours (Figure 3.12B). On the other hand, the wt MBP-VCA0681 mainly accumulated 5’-pGpG and very small amount (15%) of GMP after 20 hours (Figure 3.12A). Knock-out constructs of the first HD-domain for D75A MBP-VCA0681 showed that a) the variant is active towards c-di-GMP, b) and that it can perform its two-step hydrolysis to GMP more efficiently when compared to the wt protein (Figure 3.12B). This is in contrast with previous studies, in which D75A MBP-VCA0681 exhibited significant PDE activity, albeit it ‘stopped’ to the 5’pGpG intermediate (9). When cGAMP is employed as a substrate, D75A MBP-VCA0681 completely broke down the CDN into GMP and AMP within 2 hours (Figure 3.13). D69A MBP-SO3491 also showed a larger accumulation of GMP in 20 hour reaction compared to the wt MBP-SO3491, confirming the observation that omission of the first HD-domain facilitates the second hydrolysis step to GMP (Figure 3.12C & D).
Figure 3.12. HPLC chromatograms of the activity assays carried out with the dithionite-reduced wt, D75A MBP-VCA0681 and wt, D69A MBP-SO3491 and c-di-GMP as the substrate. All reactions were carried out for 1 hour and 20 hours, respectively. GMP elutes at circa 6 minutes, 5’-pGpG at 14 minutes and c-di-GMP at 13 minutes. A, C) 1 µM diiron wt MBP-VCA0681 and MBP-SO3491 reduced with NaDT were reacted with 50 µM c-di-GMP. B, D) 5 µM diiron D75A MBP-VCA0681 and D69A MBP-SO3491 were reacted with 50 µM c-di-MP. All reactions were carried out under O₂-free conditions in an anaerobic glovebox and terminated by heat quenching at 100 °C.
Figure 3.13. HPLC chromatograms of the activity assays carried out with the dithionite-reduced D75A MBP-VCA0681 and with c-GAMP as the substrate for 10 minutes, 40 minutes and 1 hour, respectively. GMP and AMP eluted at circa 3.5 minutes and 5.2 minutes respectively. cGAMP and 5’-pApG eluted at circa 11.6 minutes and 12.2 minutes. All reactions were carried out under O₂-free conditions in an anaerobic glovebox and terminated by heat quenching at 100 °C.

The different constructs of SO3491 were further characterized by optical spectroscopy (Figure 3.14). All proteins were in a final concentration of 100 µM except for D69A MBP-3491 (50 µM). The UV/VIS spectrum of the Fe-loaded MBP-SO3491 showed absorbances at ~ 330, 365 nm, which are characteristic of non-heme oxo/carboxylate-bridged diferric (Fe³⁺Fe³⁺) sites. The optical spectra of the wt Mn-loaded SUMO- and MBP-SO3491, as expected, lack those absorbances (because dimanganous sites in a similar ligation environment are optically transparent). The SO3491 variants were examined to determine whether knocking out of one the two HD-domains would have a detectable effect in the spectra (i.e. whether the two sites would have slightly different optical characteristics that would allow their discrimination). The absorption intensity for the knockout variants (D69A and D283A) MBP-SO3491 are in-line with
the amount of Fe cofactor estimated (which is lower than that of the wt MBP-SO3491 after normalizing considering the same protein concentration for all). On the basis of the optical spectra, we can conclude that there are diiron sites assembled in both variants, however the optical characteristics are hardly different between the two and essentially identical to those of the wt MBP-SO3491.

Figure 3.14. UV-Visible spectra of the different wt and variant SO3491 constructs in their aerobically isolated state. For all cases, the spectra are shown normalized to the protein concentration.
3.3 Bioinformatic and phylogenetic analysis

HD-GYP proteins are a relatively recent addition to the family of PDEs involved in the hydrolysis of CDNs that act as second bacterial messengers. Among these PDEs, they exhibit a large diversity both with respect to the type of cofactors they employ for catalysis and the number of domains they contain. Pathogenic bacteria encode for several HD-GYP proteins, but it is currently unclear why they contain multiple and apparently diverse HD-GYPs and what are their specific roles within the cell. To gain a deeper insight into the evolutionary and functional diversification of the different HD-GYPs, we performed a phylogenetic analysis considering all proteins, for which there is some or scant biochemical information. We aimed to establish a correlation between metal content, substrate specificity and mode of action for these HD-GYPs, with the ultimate goal to map out their multiple occurrences and specific roles within the cell.
Figure 3.15. Maximum-likelihood rooted phylogenetic tree generated with 103 sequences and including 10 sequences of the Cas3 PDE subfamily as an outgroup (31). The tree was computed with the RaxML Blackbox software assuming the LG model, a gamma model for the rate of heterogeneity, invariant sites, and empirical base frequencies. The bootstrap values showing the confidence of the nodes have been included for the major branches. The scale bar represents the number of substitutions per site. Sequences were aligned using both MAFFT and MUSCLE methods in the EMBL-EBI program. Proteins segregated in the different sub-branches are shaded with different colors.

Figure 3.16. Unrooted phylogenetic tree generated with 103 sequences and including 10 sequences of the Cas3 PDE subfamily as an outgroup (31). The tree was computed with the RaxML Blackbox software assuming the LG model, a gamma model for the rate of heterogeneity, invariant sites, and empirical base frequencies. The metal cofactor and plausible function of proteins are also listed.
We generated a phylogenetic tree with the CRISPR HD-domain PDE, Cas3, as an outgroup. The phylogeny branches out into three major clades (Figure of 3.15 & 3.16). One clade with a common ancestor includes our protein of interest, VCA0681 and SO3491. The common ground of these HD-GYP proteins is that they accumulate the linearized product 5’-pGpG. The second cluster includes HD-GYP proteins that have lost PDE function, and these are PA4108 and Bd1817 (1, 13). The structure of Bd1817 from *Bdellovibrio bacteriovorus* shows an asparagine residue coordinated in the active site, and which is purported to disrupt the ligand environment at the active site and lead to loss of PDE function (Figure 1.3) (13). The last and largest clade includes the putative tri-nuclear HD-GYP phosphodiesterases, which are typified by an extra glutamate residue important for the ligation of the third metal ion, and suggested to degrade c-di-GMP completely into GMP. This glutamate was predicted to coordinate the third metal cofactor in the trimetal binding sites and performs the two-step hydrolysis without detecting 5’-pNpN as an intermediate product (32). For these putative trinuclear PDEs, which include VCA0210, TM0186, PA4781, VCA1348 and VC2340. TM0186, VC1295 and VC2340, there is some biochemical evidence that they can hydrolyze c-di-GMP to GMP (12). VCA0210 and PA4781 contain an N-terminal phosphoreceiver (REC) domain that may activate or inhibit PDE activity through an allosteric mechanism involving its phosphorylation (19).

Although PA4781 proteins are grouped within the putative tri-nuclear HD-GYP containing proteins, it was found to exhibit a very sluggish PDE, which appears to be too slow to be functionally relevant (1, 14). We generated a sequence alignment with the PA4781 and *Pm*GH proteins (Figure 3.17). Some of the PA4781s have an extra proline instead of the extra glutamate in the active site, which may result in the loss of PDE activity and ability to coordinate the third
metal. The other PA4781s have lost the conserved double HH sequence; this missing HHE sequence may be another plausible reason for the loss or very slow PDE activity.

Figure 3.17. Sequence alignment of PA4781 and tri-nuclear HD-GYP Proteins. The conserved extra glutamates and HHE sequences are highlighted in green box with an arrow pointing towards them.
Chapter 4

Discussion

We have biochemically and structurally characterized the VCA0681 homolog, SO3491 from the model γ-proteobacterium *Shewanella oneidensis*. We established that SO3491 is a true functional homolog of VCA0681 from *Vibrio cholera*. SO3491 showed similar metal dependence and comparable activities with VCA0681. The PDE activity of VCA0681 is strongly Fe-dependent, but can also be sustained by Mn (albeit less efficiently). The Fe-loaded VCA0681 is highly active with c-di-GMP and is the first reported enzyme that can degrade the novel cyclic dinucleotide c-GAMP. SO3491 is also as active as VCA0681 with both c-di-GMP and c-GAMP. SO3491 purified with different metal cofactors (Fe and Mn, respectively) allowed us to examine the dependence of its catalytic activity on the chemical nature of the cofactor and its ability to hydrolyze different substrates. Mn-loaded SO3491 demonstrated hydrolysis of c-di-GMP to 5’-pGpG, although very slow compared to the Fe-loaded SO3491. Phosphodiesterase activities of VCA0681 and SO3491 are limited to c-di-GMP and c-GAMP substrates, and the product of their reactions is the one-step hydrolyzed 5’-pGpG and 5’-pApG linearized dinucleotides, respectively.

VCA0681 and SO3491 harbor two diiron sites, among which one has the HD-GYP motif and the other one contains only the HD motif. The diiron binding sites of VCA0681 and SO3491
can attain three redox states, the diferric, mixed-valent and the diferrous forms. The Fe\textsuperscript{Fe\textsuperscript{II}} was accumulated to a certain extent by reduction with small excess amount of sodium ascorbate. The Fe\textsuperscript{Fe\textsuperscript{II}} was attained in homogeneity (100\%) by reduction with excess amount of sodium dithionite. However, quantification on the basis of the Mössbauer and EPR spectra, Fe\textsuperscript{Fe\textsuperscript{II}} accumulates only in small percentage (16\%) in SO3491 and Fe\textsuperscript{Fe\textsuperscript{II}} accumulation (19\%) cannot be avoided in the ascorbate reduced protein. Our results indicate that Fe\textsuperscript{Fe\textsuperscript{II}} of SO3491 either does not contribute detectably to the activity or is much slower than the Fe\textsuperscript{Fe\textsuperscript{II}} in hydrolyzing c-di-GMP. The catalytic form Fe\textsuperscript{Fe\textsuperscript{II}} is quite sensitive against O\textsubscript{2} oxidative inactivation and can react with O\textsubscript{2} with a second-order rate constant of 20 mM\textsuperscript{-1}s\textsuperscript{-1}.

Steady-state kinetic assays were carried out to determine the catalytic efficiency of SO3491. Wt native (‘untagged’) SO3491 hydrolyzes c-di-GMP with $k_{cat}$ and $K_M$ of 0.15 s\textsuperscript{-1} and 3.1 µM respectively. The cellular levels of c-di-GMP in V. cholerae are within the range of 0.9 to 20 µM (24). SO3491 is enzymatically quite proficient when compared on the basis of its catalytic efficiency to other HD-GYP PDEs (Table 2), to hydrolyze c-di-GMP and thus switch the bacterial behavior to non-pathogenic one. SO3491 hydrolyzes c-GAMP with a $k_{cat}$ of $5.3 \times 10^{-3}$ s\textsuperscript{-1} and $K_M$ of 55 µM, which is much slower than c-di-GMP. Currently there is no other report for the catalytic proficiency of these enzymes with c-GAMP as a substrate, not the cellular levels of c-GAMP have been to our knowledge estimated.

Variants of VCA0681 and SO3491, in which one of the HD dimetal sites was knocked out, were assayed to explore each site’s unique role. We found that the C-terminal HD domain (HD\textsubscript{\textsuperscript{289}} and HD\textsubscript{\textsuperscript{283}}, in VCA0681 and SO3491, respectively) is essential for PDE activity. The absence of this HD-domain completely abolishes the PDE activities of both VCA0681 and SO3491. On the other hand, knocking out the N-terminal HD domain does not eliminate PDE activity. In contrast,
it allows for a two-step hydrolysis to take place, most likely through a structural rearrangement or flexibility imposed on the C-terminal ‘active’ HD-GYP domain, which is now competent breakdown the cyclic di-nucleotide completely to the individual nucleotides. While currently the exact functional purpose of the N-terminal HD domain is not precisely known, we propose that it serves mainly a structural role and that it can divert the reaction outcome via allosteric or conformational changes imposed on the C-terminal domain. Overall it appears to serve a ‘stop’ function, so that the reaction can be halted to the 5’pGpG product, which may be relevant for the cell, especially because 5’pGpG has been proposed to serve a similar function to c-di-GMP (i.e. a second bacterial messaging molecule).

The bioinformatic and phylogenetic analysis informed us the biochemical and structural diversion of HD-GYP proteins, for which only a few representatives have been characterized. HD-GYP proteins appear to have segregated into three main groups on the basis of their metal cofactors, their respective specific activities and their reaction outcome. One phylogenetic cluster of HD-GYP proteins contains representatives containing putative tri-nuclear metal binding sites and proposed to afford a two-step hydrolysis that can break down cyclic dinucleotides completely into the individual nucleotides without obvious accumulation of the linearized intermediates. Another group of HD-GYP phosphodiesterases, including our proteins of interest VCA0681 and SO3491, on the other hand accumulate mainly the linearized 5’-pNpN products after first-step hydrolysis, which may be related to their different roles within the cell and their expression under different environmental metal ion concentrations. The last clade of the phylogenetic tree contains HD-GYP proteins that are enzymatically inactive. These proteins can cooperate with different metals at their active sites, suggesting that the metal binding sites may play a structural role more than a catalytic role (14). These have been shown to bind c-di-GMP, which would suggest that their likely evolved
function is to act as CDN sensors rather than CDN degraders. Overall our phylogenetic analysis suggests that there are three main classes of HD-GYP proteins, which are further branching out to smaller clades that demonstrate slightly different activities, metal ion content and substrate specificities. This divergence is a most likely outcome of an evolutionary diversification of HD-GYPs as a response to different environmental cues. We believe that our analysis can serve as a functional map predictor that allows portraying the somewhat different roles of these proteins in the cell, and opens up the stage for our better understanding for the regulation of important second bacterial messengers, such as c-di-GMP.
Chapter 5

Future Directions and Experiments

We hypothesized that the mutation of the N-terminal HD domain strongly destabilizes the protein. This hypothesis need to be verified by thermoshift denaturation assays to obtain the melting curves of the D283A MBP-SO3491. The site differentiation on the Mössbauer spectra could either correspond to the different HD-domains of the protein or to a mixed population of active site molecules with a slight deviant ligation or the presence of a small ligand bound to the active site. In order to clarify the reason caused this differentiation, Mössbauer spectra on the MBP-D69A should be obtained, where only one HD-domain is presented. It is also important to study the kinetic parameters of D69A MBP-SO3491, wt MBP-VCA0681 and D75A MBP-VCA068, which can give more knowledge on the catalytic proficiency of distinct HD-domain phosphodiesterases.
Bibliography

8. Sun L, Wu J, Du F, Chen X, Chen ZJ. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. Science 2013; 339:786-791


