A Novel Capture Compound for the Identification and Analysis of Cyclic di-GMP Binding Proteins

Jutta Nesper[1)], Alberto Reinders[1)], Timo Glatter*[1)], Alexander Schmidt*, and Urs Jenal

Affiliations:

Biozentrum of the University of Basel, Klingelbergstrasse 50, CH-4054 Basel, Switzerland

*Proteomics Core Facility, Biozentrum of the University of Basel, Klingelbergstrasse 50, CH-4054 Basel, Switzerland

1) These authors contributed equally to this work

For correspondence: urs.jenal@unibas.ch

Keywords: *Caulobacter crescentus*; *Pseudomonas aeruginosa*; *Salmonella enterica serovar typhimurium*; c-di-GMP effector; EAL GGDEF; Capture Compound mass spectrometry

Abstract

The second messenger cyclic di-GMP is a near-ubiquitous signaling molecule that globally alters bacterial cell physiology to promote biofilm formation and community behavior. Much progress was made in recent years towards the identification and characterization of diguanylate cyclases and phosphodiersterases, enzymes involved in the synthesis and degradation of this signaling compound. In contrast, our knowledge of the nature and mechanistic details of c-di-GMP effector proteins lags behind, primarily because effective tools for their specific enrichment and rapid analysis are missing. In this report we demonstrate that a novel tri-functional c-di-GMP-specific Capture Compound (cdG-CC) can be effectively used to identify and validate c-di-GMP binding proteins. The cdG-CC was able to specifically and efficiently pull-down *bona fide* c-di-GMP effector proteins. Furthermore, in combination with mass spectrometry (CCMS), this technology robustly identified a substantial fraction of the known c-di-GMP signaling components directly from cell extracts of different model organisms. Finally, we applied the CCMS technique to profile c-di-GMP binding proteins of *Salmonella enterica serovar typhimurium*. Our studies establish CCMS as a powerful and versatile tool to identify and analyze components of the cellular c-di-GMP pathway in a wide range of different organisms.

Cyclic di-GMP is a ubiquitous second messenger regulating growth and behavior of a wide range of gram-positive and gram-negative bacteria. In particular, c-di-GMP mediates the switch between single cell behavior and a community life style called biofilm, which is often associated with chronic infections of bacterial pathogens [1]. Major components of the regulatory network are the GGDEF and EAL domains that are widespread in bacteria and catalyze c-di-GMP synthesis and degradation, respectively [2]. The list of cellular processes controlled by c-di-GMP is ever-increasing and includes the biosynthesis and secretion of surface adhesins and exopolysacharide (EPS) matrix components, different forms of cellular motility, long-term survival and response to environmental stress, synthesis of secondary metabolites, regulated proteolysis and cell cycle progression, delivery of anti-bacterial toxins, intracellular growth and the production of virulence factors in a range of animal and plant pathogens [2-4]. Despite of this wide variety of cellular functions that are modulated by c-di-GMP, the list of effector proteins has remained relatively sparse [2-4]. These include PilZ, small switch-like domains that undergo conformational change upon binding c-di-GMP. In addition, several members of the CRP/FNR and response regulator superfamilies of transcription regulators were shown to specifically bind c-di-GMP. Finally, a subgroup of GGDEF and EAL domains was recognized as c-di-GMP effector proteins adopting their novel functionality through the combined loss of catalytic activity and exploitation of their allosteric and active site binding pockets, respectively. While most known effector proteins were discovered through an “educated guess” approach that was based on their functional linkage to c-di-GMP mediated cellular processes, unbiased screening for novel effectors was hampered primarily by the lack of reliable and effective biochemical tools for their enrichment and isolation. Only one global recent study used a chemical proteomics approach to identify c-di-GMP binding proteins in *P. aeruginosa* [5]. Here we introduce a novel tri-functional capture molecule (cdG-CC) as an effective tool to identify specific c-di-GMP binding proteins directly from a complex mixture of macromolecules. The compound is based on a chemical scaffold harboring specificity, reactivity, and sorting properties (Fig. 1A) [6, 7].

To evaluate the potential of this compound for the specific enrichment and isolation of c-di-GMP binding proteins we first tested if the cdG-CC is able to pull down known

c-di-GMP effector molecules directly from a solution with purified protein. We used the well-studied PilZ domain protein DgrA from *C. crescentus* [8]. The binding affinity of wild-type DgrA for c-di-GMP is in the low nanomolar range. In contrast, the binding mutant DgrAW75A has a Kd of 6.2 μM and a DgrA mutant lacking coordinating Arg residues in the N-terminus (RR11AA) fails to bind c-di-GMP completely [8]. 0.5 μM purified wild-type and mutant hexahistidine-tagged proteins (Fig. 1B) were incubated with 10 μM cdG-CC, UV cross-linked for 4 min and pulled out with streptavidin coated magnetic beads. While DgrA was readily captured in the absence of any other ligand, the addition of a 100x excess of c-di-GMP abolished cdG-CC almost entirely (Fig. 1B). In contrast, an equal molar excess of GTP did not interfere with the cdG-CC pull down (Fig. 1B) arguing that cdG-CC binding to DgrA is highly specific. In agreement with this, the cdG-CC only inefficiently pulled-down the DgrAW75A mutant form and completely failed to bind the DgrARR11AA mutant (Fig. 1B). These data are fully consistent with the binding affinities of the different DgrA forms for c-di-GMP and strongly argue that cdG-CC is able to specifically capture DgrA. Moreover, these experiments demonstrate that the cdG-CC is a valuable diagnostic tool to verify candidate c-di-GMP binding proteins.

These results encouraged us to probe if the capture compound can be applied for the selective enrichment of c-di-GMP binding proteins from a more complex mixture of proteins. To test this, we captured soluble c-di-GMP binding proteins from cell extracts and probed immunoblots with PopA specific antibodies. PopA is a GGDEF effector protein that regulates cell cycle progression in *C. crescentus* in response to a cellular upshift of c-di-GMP during the G1-S phase transition [9, 10]. To bind c-di-GMP PopA utilizes a conserved and well-defined binding pocket, which, in related catalytic GGDEF domains, is used as an allosteric I-site for product inhibition of the diguanylate cyclase (DGC) activity [11, 12]. As shown in Fig. 1C, PopA with a known Kd for c-di-GMP of 2.5 μM, was readily captured from *C. crescentus* cell lysates using 3 μM cdG-CC and a total of 400 μg soluble protein. Pull down of PopA was inhibited in the presence of a large excess of c-di-GMP (1mM), while GTP did not interfere with cdG-CC binding (Fig. 1C), arguing that the cdG-CC interaction with PopA is highly specific. Likewise, no PopA was bound to the cdG-CC when using a strain expressing a PopA mutant that lacks the highly conserved Arg residue of the canonical RxxD I-

site binding motif (R357G) [9] (Fig. 1C). This indicated that the cdG-CC can enrich c-di-GMP binding proteins directly from whole cell extracts in a highly specific manner and that this compound is suited for a global isolation procedure of c-di-GMP binding proteins. Therefore we combined capture experiments with the analysis of isolated proteins by LC-MS/MS (CCMS, [6, 7]). When applying 10, 5 or 2.5 μM cdG-CC with 400 μg of soluble *C. crescentus* proteins, nine of eleven proteins predicted to contain either a PilZ, GGDEF or EAL domain (Table 1A and Supplementary Table S1A) were significantly enriched as compared to the competition control based on the spectral counts of the identified peptides. In addition to the analysis of soluble proteins we also aimed at evaluating the efficiency of CCMS for the enrichment of c-di-GMP binding proteins from membrane fractions. Although the numbers of spectral counts were lower as compared to proteins from the soluble fraction, three of the five integral membrane proteins predicted to bind c-di-GMP were identified when 400 μg DDM solubilized membrane proteins and 10 μM cdG-CC were used for CCMS (Table 1A and Supplementary Table S1A). Only four of the known components of the *C. crescentus* c-di-GMP network were not identified by CCMS. Two of these are integral membrane proteins with several predicted membrane spanning domains in their N-terminal regions (CC0740, CC0896). It is possible that they were not solubilzed by the detergent used or not detected by LC-MS/MS. Another possibility is that they are not expressed, as it might be the case for CC3094 and CC3148.

To expand these studies to a different organism, we also tried to isolate known c-di-GMP binding proteins from *P. aeruginosa* PA01. With over 40 GGDEF and EAL domain and eight PilZ domain proteins annotated, the complexity of the c-di-GMP signaling network in this organism is much higher than in *C. crescentus*. Using 10 μM, 7.5 μM, 5 μM, 2.5 μM and 1.25 μM cdG-CC and 350 μg of soluble protein extract, several of these proteins were unambiguously identified by CCMS (Table 1B and Supplementary Table S1B). This includes four PilZ domain proteins, four GGDEF and two composite GGDEF-EAL domain proteins. The fraction of the *P. aeruginosa* proteins predicted to bind c-di-GMP that were isolated by CCMS is substantially lower (10 of 28 predicted soluble proteins) as compared to *C. crescentus* but is comparable to a recent study using sepharose-coupled c-di-GMP [5]. The

discrepancy between *C. crescentus* and *P. aeruginosa* might be related to the fact that many of these proteins are engaged in the transition between a motile, planktonic and a sessile, surface attached life-style. Since *C. crescentus* has integrated this developmental transition into its reproductive cycle [13], most of the components of the c-di-GMP network orchestrating this switch, need to be expressed in cells growing in liquid cultures [10]. In contrast, *P. aeruginosa* adapts its expression program upon surface colonization in response to environmental cues [14]. It is thus possible that several components of the *P. aeruginosa* core machinery required to adapt to surface communities are simply not expressed in logarithmically growing cultures in liquid media. To test this, it would be interesting to repeat theses experiments with cell extracts harvested from surface grown *P. aeruginosa* communities.

To evaluate whether cdG-CC based CCMS can be used to profile an entire cellular network of c-di-GMP binding proteins we used *S. typhimurium*. This organism was chosen to perform an unbiased CCMS experiment because its c-di-GMP network appears to be of lower complexity as compared to other bacterial species [15]. In contrast to the CCMS experiments with *C. crescentus* and *P. aeruginosa*, in which spectral counts were extracted for known c-di-GMP binders, we tested whether CCMS is capable to enrich for c-di-GMP binding proteins upon MS1 label-free quantification. CCMS experiments were performed in triplicates with 350 μg soluble whole cell proteins using 7.5 μM cdG-CC and competition controls with an excess of c-di-GMP (1 mM) (Supplementary Table S2A). Following mass spectrometry analysis and label-free quantification (Supplementary Table S2B) significant differences in protein enrichment between cdG-CC experiments and control samples with competing c-di-GMP were visualized in a volcanoplot. The graph shows a significant enrichment (>2 fold) of 36 proteins as compared to the control with a q-value <0.05 (Supplementary Table S2C and Fig. 2). Among the enriched proteins is the PilZ domain protein YcgR (Fig. 2) [16]. Many of the identified components that were not previously associated with the c-di-GMP network were metabolic proteins and proteins involved in fatty acid and LPS biosynthesis (Fig. 2). Such proteins are of great interest in the light of switching between a virulent planktonic and a surface attached persistent lifestyle. However, since none of these proteins show homology

to known c-di-GMP binding proteins, they first need to be validated to specifically bind c-di-GMP and be part of the c-di-GMP signaling network. In addition also some background proteins are enriched. They are mostly binders of other nucleotides or proteins that are known to be highly abundant, such as five tRNA-related proteins or the chaperone GroEL (Supplementary Table S2C).

Acknowledgements

We thank Samuel Steiner for help in optimization of Capture Compound binding assays. This work was supported by the Swiss National Science Foundation (SNF) Sinergia grant CRSII3_127433.

References:

[1] Furukawa S, Kuchma SL, O'Toole GA. Keeping their options open: Acute versus persistent infections. J Bacteriol. 2006;188:1211-7.

[2] Schirmer T, Jenal U. Structural and mechanistic determinants of c-di-GMP signalling. Nat Rev Microbiol. 2009;7:724-35.

[3] Hengge R. Cyclic-di-GMP reaches out into the bacterial RNA world. Sci Signal. 2010;3:pe44.

[4] Sondermann H, Shikuma NJ, Yildiz FH. You've come a long way: C-di-GMP signaling. Curr Opin Microbiol. 2012.

[5] Duvel J, Bertinetti D, Moller S, Schwede F, Morr M, Wissing J, et al. A chemical proteomics approach to identify c-di-GMP binding proteins in *Pseudomonas aeruginosa*. J Microbiol Methods. 2012;88:229-36.

[6] Koster H, Little DP, Luan P, Muller R, Siddiqi SM, Marappan S, et al. Capture compound mass spectrometry: A technology for the investigation of small molecule protein interactions. Assay Drug Dev Technol. 2007;5:381-90.

[7] Lenz T, Fischer JJ, Dreger M. Probing small molecule-protein interactions: A new perspective for functional proteomics. J Proteomics. 2011.

[8] Christen M, Christen B, Allan MG, Folcher M, Jeno P, Grzesiek S, et al. DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*. Proc Natl Acad Sci U S A. 2007;104:4112-7.

[9] Duerig A, Abel S, Folcher M, Nicollier M, Schwede T, Amiot N, et al. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. Genes Dev. 2009;23:93-104.

[10] Abel S, Chien P, Wassmann P, Schirmer T, Kaever V, Laub MT, et al. Regulatory cohesion of cell cycle and cell differentiation through interlinked phosphorylation and second messenger networks. Mol Cell. 2011;43:550-60.

[11] Christen B, Christen M, Paul R, Schmid F, Folcher M, Jenoe P, et al. Allosteric control of cyclic di-GMP signaling. J Biol Chem. 2006;281:32015-24.

[12] Wassmann P, Chan C, Paul R, Beck A, Heerklotz H, Jenal U, et al. Structure of BeF3- -modified response regulator PleD: Implications for diguanylate cyclase activation, catalysis, and feedback inhibition. Structure. 2007;15:915-27.

[13] Curtis PD, Brun YV. Getting in the Loop: Regulation of development in *Caulobacter crescentus*. Microbiology and Molecular Biology Reviews. 2010;74:13-41.

[14] Kuchma SL, O'Toole GA. Surface-induced and biofilm-induced changes in gene expression. Curr Opin Biotechnol. 2000;11:429-33.

[15] Solano C, Garcia B, Latasa C, Toledo-Arana A, Zorraquino V, Valle J, et al. Genetic reductionist approach for dissecting individual roles of GGDEF proteins within the c-di-GMP signaling network in *Salmonella*. Proc Natl Acad Sci U S A. 2009;106:7997-8002.

[16] Ryjenkov DA, Simm R, Romling U, Gomelsky M. The PilZ domain is a receptor for the second messenger c-di-GMP: The PilZ domain protein YcgR controls motility in enterobacteria. J Biol Chem. 2006;281:30310-4.

Figure legends:

Fig. 1 - The cdG-CC can specifically pull down bona fide c-di-GMP binding proteins from purified and from crude cell extracts. A) Chemical structure of cdG-CC. C-di-GMP provides the selectivity for capturing proteins. The photo-reactivity group allows irreversible nitren formation between the compound and the captured protein. Biotin as sorting function allows the isolation of the compound by binding to streptavidin coated magnetic beads. B) Immunoblot of purified and captured DgrA (wt), DgrA W75A and DgrA RR11AA. In all capture experiments 10 μM cdG-CC was present, in competition experiments the proteins were preincubated with a 100 x excess of c-di-GMP (cdG) or GTP. All proteins were His-tagged and detected using anti-His antibodies. C) Immunoblot of PopA. PopA (wt) or the PopA I-site mutant R357G was expressed in NA1000ΔpopA (lanes marked with 'proteins'; note that this I-site mutant is less abundant in the cell). PopA was captured in the presence of 3 μM cdG-CC. Addition of 1 mM c-di-GMP but not 1 mM GTP prevented the binding to the cdG-CC. In contrast the PopA I-site mutant could not be captured at all. PopA was detected using anti-PopA antibodies.

Fig. 2 - Volcanoplot based visualization of proteins significantly enriched by CCMS of *S. typhimurium*. Following capturing, LC-MS/MS analysis and label-free quantification, log2-intensity ratio of all detected peptide features between capturing and competition experiment were calculated and plotted versus values derived from significance analysis. Proteins within the significance thresholds for q-values <0.05 and intensity ratios >2-fold are indicated in a box. Experiments in triplicate were performed in the presence of 7.5 μM cdG-CC and with 1 mM c-di-GMP added to the competition reactions.

Table 1 - Identified known c-di-GMP binding proteins.

A) *C. cresentus*

Protein Name	**ID**	**Domain architecture**	**CCMS experiment/CCMS competition[1)] No spectral counts of identified peptides**		
Soluble fraction: Experiment [2)]			1	2	3
PopA	CCNA_01918	GGDEF	20/1	25/3	15/5
PleD	CCNA_02546	Rec-Rec-GGDEF	18/0	21/2	13/5
DgcA	CCNA_03394	GGDEF	5/0	3/0	1/0
DgcB	CCNA_01926	GGDEF	12/1	16/3	10/2
CC0655	CCNA_00692	PAC-GGDEF-EAL	8/0	18/0	19/0
CC1086	CCNA_01140	PAS-EAL	8/0	5/0	4/0
PdeA	CCNA_03507	GGDEF-EAL	3/0	5/0	6/0
DgrA	CCNA_01671	PilZ	5/0	6/3	6/1
DgrB	CCNA_03268	PilZ		2/0	
Membrane fraction: Experiment[3)]					
TipF	CCNA_00747	EAL	3/0	23/13	
PdeB	CCNA_00089	3x(MHYT)-PAS-GGDEF-EAL	3/0	6/0	
CC0857	CCNA_00900	CHASE4-GGDEF-EAL		2/0	

1) All competition experiments were performed in the presence of 1 mM c-di-GMP.
2) 3 independent experiments are indicated using 10, 5 or 2.5 μM cdG-CC respectively.
3) 2 independent experiments are indicated using 10 μM cdG-CC.

B) Soluble proteins of *P. aeruginosa*

Protein Name	**ID**	**Domain architecture**	**CCMS experiment/CCMS competition[1)] No of spectral counts of identified peptides**				
Experiment No[4)]			1	2	3	4	5
FimX	PA4959	PAS-GDSIF-EVL	9/0	10/0	9/0	8/0	7/0
	PA3353	PilZ	5/0	6/0	6/0	10/0	7/0
WspR	PA3702	Rec-GGEEF	2/0	2/0	2/0		
	PA0012	PilZ			2/0	1/0	
	PA0169	GGEEF	2/0	2/0	2/0		
	PA4843	Rec-Rec-GGEEF		2/0	2/0	5/0	3/0
	PA2989	PilZ	1/0	1/0	1/0	1/0	1/0
	PA0290	PAS-Rec-GGEEF	1/0			2/0	2/0
	PA2567	GAF-SPTRF-EAL		1/0	1/0		1/0
	PA4608	PilZ	1/0	2/0	2/0	2/0	1/0

4) Eperiment 1 was performed with 10 µM cdG-CC, experiment 2 with 7.5 µM cdG-CC, experiment 3 with 2.5 µM cdG-CC, experiment 4 with 2.5 µM cdG-CC and experiment 5 with 1.25 µM cdG-CC.

Fig. 1
Click here to download high resolution image

Figure 1:

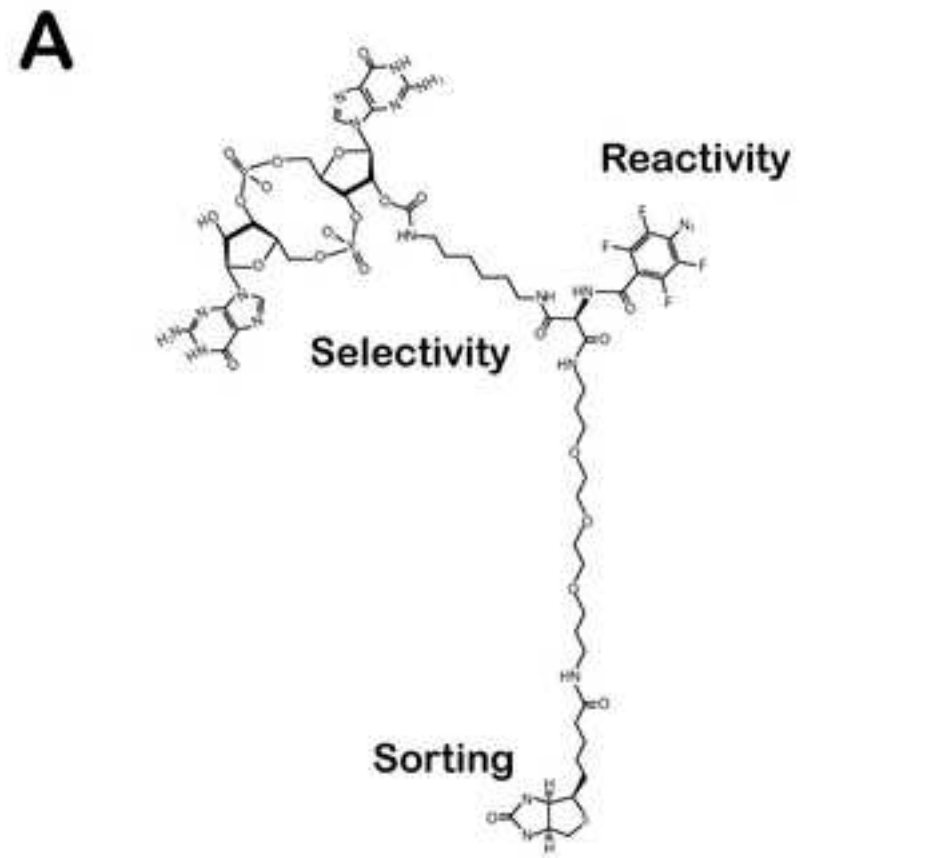
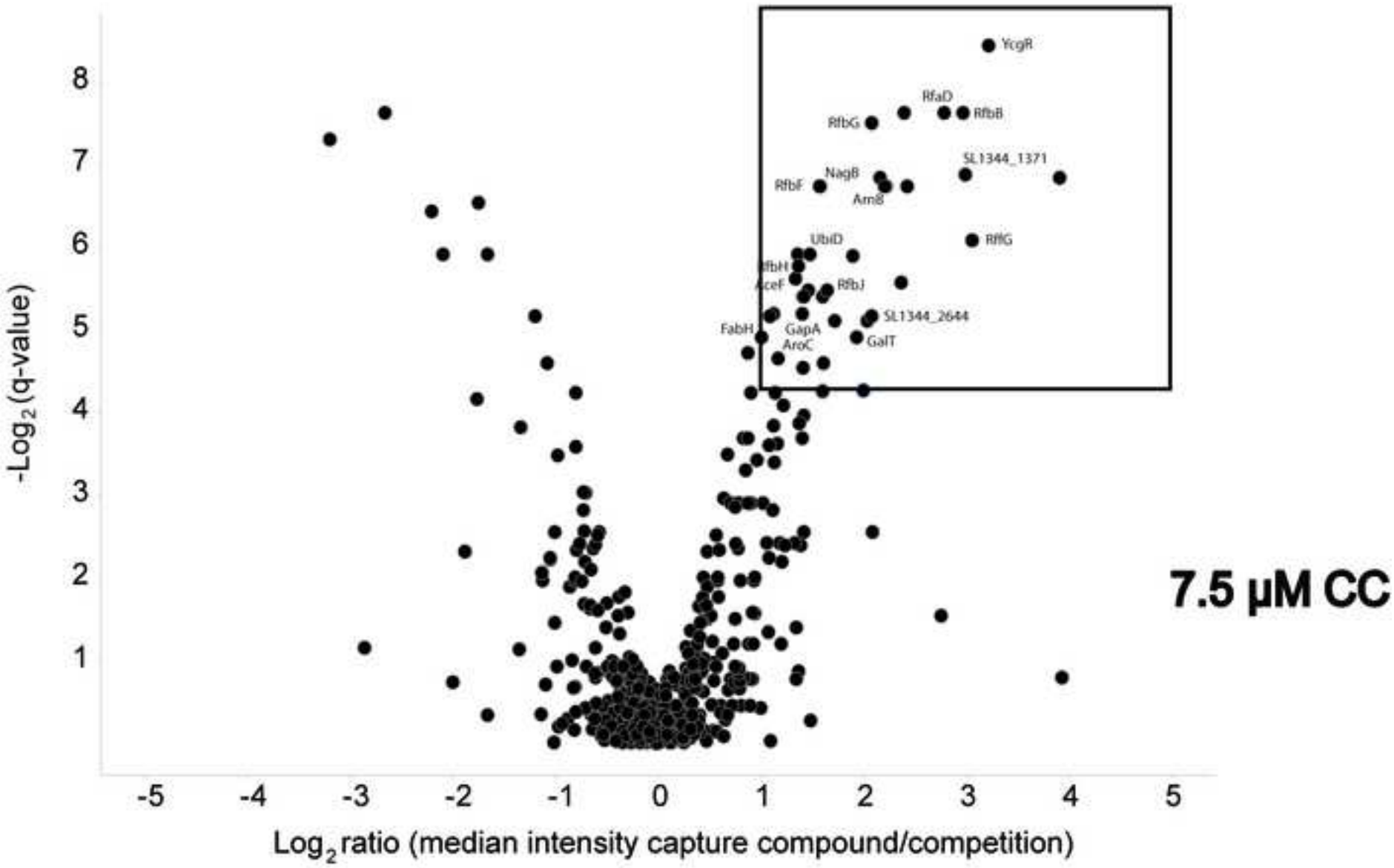


Fig. 2

Click here to download high resolution image

Figure 2:



Materials & Methods

Click here to download Supplementary material: SupplementaryMaMe.docx

Table S1

Click here to download Supplementary material: SupplementaryTable S1.xlsx

Table S2

Click here to download Supplementary material: SuplementraryTable S2.xlsx