Signal Transduction in Membrane-Bound Adenylate Cyclases

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Jens Baßler

aus Esslingen am Neckar

Tübingen

2018

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	16.07.2018
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Doron Rapaport
2. Berichterstatter:	Prof. Dr. Andrei Lupas
3. Berichterstatter:	apl. Prof. Dr. Rainer Merkl

Abstract

Class III adenylate cyclases (ACs) are widespread signaling proteins, which translate diverse intracellular and extracellular stimuli into a uniform intracellular signal. They are typically composed of an N-terminal array of sensor domains and transducers, followed C-terminally by a catalytic domain, which, after dimerization, generates the second messenger cyclic adenosine monophosphate (cAMP). Many of the N-terminal domains are also found in other signaling proteins and can frequently be recombined between them.

This work bioinformatically investigates the architectural and evolutionary principles that enable the productive interaction of a great diversity of upstream regulatory domains with the conserved AC catalytic domain. As part of this process, we have identified the novel cyclase transducer element (CTE), a pivotal hinge on the N-terminus of the AC catalytic domain. The element appears to convert unspecific signals moving along the coiled-coil backbone into specific conformational changes that determine AC activity. This suggests communication along a dimeric coiled coil as the structural rationale for the architectural similarities between many families of signaling proteins.

Further, we bioinformatically classify the various six-helical transmembrane (6TM) domains observed in many bacterial and eukaryotic ACs. Recently, experimental results have indicated that these domains could function as regulatory receptors binding hydrophobic ligands inside the membrane. Our classification and the presence of a CTE in many 6TM ACs strongly support this hypothesis, concluding that some or all 6TM domains are regulatory receptors for as-yet-unknown ligands. Since mammalian ACs are important downstream messengers of G protein-coupled receptors, these findings suggest AC 6TM domains could be drug targets of possibly great pharmacological relevance.

Zusammenfassung

Klasse III Adenylatcyclasen (ACn) sind weitverbreitete Signaltransduktionsproteine, die verschiedenste intra- und extrazelluläre Reize in ein einheitliches intrazelluläres Signal umwandeln. Zumeist bestehen diese Proteine aus einer Reihe N-terminaler Domänen zur Wahrnehmung und Weiterleitung regulatorischer Reize, gefolgt von der C-terminalen katalytischen Domäne, welche, nach Dimerisierung, den Sekundärbotenstoff cyclisches Adenosinmonophosphat (cAMP) produziert. Viele der N-terminalen Domänen finden sich zudem auch in weiteren Signaltransduktionsproteinen, zwischen welchen sie sich häufig austauschen lassen.

Diese Arbeit untersucht mit bioinformatischen Mitteln die strukturellen und evolutionären Grundlagen welche die produktive Interaktion zwischen dieser großen Vielfalt vorgeschalteter regulatorischer Domänen und der konservierten katalytischen Domäne ermöglichen. Im Zuge dieser Untersuchung identifizierten wir das neuartige Cyclase Transducer Element (CTE), eine funktional entscheidende Gelenkstruktur am N-terminus der katalytischen Domäne. Dieses Element wirkt mutmaßlich als Konverter unspezifischer Signale aus dem coiled-coil Rückgrat vieler ACn zu einer spezifischen Konformationänderung, welche direkt die Aktivität der AC beeinflusst. Dies deutet zudem darauf hin, dass die Kommunikation entlang einer dimeren Coiled Coil das gemeinsame Prinzip hinter den oftmals ähnlichen Domänenarchitekturen vieler verschiedener Signalstransduktionsproteine sein könnte.

Des Weiteren klassifizierten wir die verschiedenen sechshelikalen Transmembrandomänen (6TM) vieler bakterieller und eukaryotischer ACn. Kürzlich veröffentlichte Ergebnisse deuten darauf hin, dass diese Domänen regulatorische Rezeptoren für hydrophobe Liganden innerhalb der Membran sein könnten. Die Ergebnisse dieser Klassifizierung sowie die Identifizierung von CTEn in vielen ACn mit 6TM Domänen unterstützen diese Hypothese. Wir schließen hieraus, dass ei-

nige oder alle dieser Domänen Rezeptoren für derzeit noch unbekannte Liganden darstellen. Da die membrangebundenen ACn von Säugetieren Teil der Signaltransduktionskette von G-Protein gekoppelten Rezeptoren sind, empfiehlt dies ACn als attraktive Wirkziele für die Entwicklung zukünftiger Medikamente.

My religion, such as it is, is that we are shaped by the universe to be its consciousness. We tell the universe what it is. In my religion, the building of a telescope is the building of a cathedral. (...) I would much rather be a rising ape than a fallen angel.

Sir Terry Pratchett

Acknowledgement

N-terminally, at the site of stimulatory input, I would like to express my gratitude to Prof. Dr. Andrei Lupas for his supervision, his support, his continued trust, and for teaching me to prefer sequences over papers and structures over figures.

I want to thank Prof. Dr. Joachim Schultz for his supervision, for introducing me to adenylate cyclases, and for many long and insightful discussions about scientific and non-scientific topics.

I am thankful to Prof. Dr. Doron Rapaport, Prof. Dr. Thilo Stehle and apl. Prof. Dr. Rainer Merkl for being part of my examination committee, for their helpful comments, and for always fast and uncomplicated contact.

I was extremely privileged to work on this project together with many talented colleagues: Dr. Reinhard Albrecht, Dr. Stephanie Beltz, Manuel Finkbeiner, Julia Grischin, Ina Meuskens, Anita Schultz, Anubha Seth, and Dr. Miriam Ziegler - thank you so much!

I would like to thank Janet Giehl, Philipp Kärgel, Dr. Joana Pereira, and Laura Weidmann for their critical reading and helpful comments on this thesis.

And C-terminally, where the output is ultimately decided, I want to thank all my family and friends, to whom I wish to express my gratitude in deeds rather than in words. Without you, this had not been possible.

Contents

List of Abbreviations 1

1 Introduction 5

1.1 Cellular signal transduction 5

1.2 ACs produce the second messenger cAMP 7

1.3 Bacterial ACs are modular multidomain proteins 11

1.4 Eukaryotic mACs resemble bacterial 6TM ACs 14

1.5 AC 6TM domains could be regulatory receptors 16

1.6 Chimeras between bacterial quorum sensors and ACs 17

2 Objectives 19

3 Results and Discussion 21

3.1 Classification of AC 6TM domains 21

3.2 The cyclase transducer element 25

3.3 Comprehensive analysis of AC domain architectures 32

4 Conclusion and Outlook 35

References 37

Appendix 47

A Manuscript Beltz et al. (2016) eLIFE 49

B Manuscript Ziegler et al. (2017) FEBS J 69

C Manuscript Bassler et al. (2018) Cellular Signalling 85

D Curriculum vitae 96

List of Abbreviations

Note: In the interest of readability, I will treat some acronyms, including domain names and chemicals, as proper names in the text when the long form is not required for context. The full definitions for all acronyms in this dissertation are given below.

Abbreviations

AC	adenylate cyclase
AMP	adenosine monophosphate
ATP	adenosine triphosphate
BLAST	basic local alignment search tool (software)
c-di-GMP	cyclic di-guanosine monophosphate
CAI-1	*Vibrio cholerae* autoinducer 1
cAMP	cyclic adenosine monophosphate
CAP	catabolite activator protein
cGMP	cyclic guanosine monophosphate
CLANS	clustered analysis of sequences (software)
CRP	cAMP receptor protein
di-GC	diguanylate cyclase
FhlA	formate hydrogen-lyase transcriptional activator
GC	guanylate cyclase
GEF	guanine nucleotide exchange factor
GPCR	G protein-coupled receptor
Gsα	G protein subunit alpha
GTP	guanosine triphosphate
HK	histidine kinase

HMM	hidden Markov model
mAC	membrane-bound AC
MCP	methylated chemotaxis protein
NC	nucleotide cyclase
NTP	nucleotide triphosphate
PDB	protein data bank
PDE	phosphodiesterase
PKA	protein kinase A
RR	response regulator
sAC	soluble AC
TCST	two-component signal transduction

Domain Names

6TM	six-helical transmembrane domain
BLUF	input domain sensing blue light using FAD
C1b	conserved linker between the subunits of mAC pseudo-heterodimers
CACHE	input domain found in calcium channel and chemotaxis receptor proteins
CHASE2	cyclase/histidine kinase-associated sensing extracellular domain type 2
CTE	cyclase transducer element
DHp	dimerization and histidine phosphotransfer domain
FHA	forkhead-associated domain
GAF	input domain found in cGMP-specific phosphodiesterases, adenylate cyclases, and FhlA
HAMP	transducer domain found in histidine kinases, adenylate cyclases, methylated chemoreceptors, and phosphatases
PAS	input domain found in period circadian protein, aryl hydrocarbon receptor nuclear translocator protein, and single-minded protein
REC	phosphorylation receiver domain
SMBD	small molecule-binding domain
TPR	tetratrico peptide repeat-like domain

1 Introduction

1.1 Cellular signal transduction

All life is cellular, separated from its environment by the plasma membrane. While indispensable for the upkeep of most biochemical processes, this barrier also separates the cell from vital information about its surroundings, for example the availability of nutrients, osmotic conditions, or the presence of signal molecules from other cells. Hence, cells enrich their membranes with dedicated proteins that transduce relevant extracellular stimuli across the membrane into the cytosol. The individual set of signaling proteins defines an organism's capabilities to react to its environment, and is thus tightly linked to its adaptation to a particular ecological niche. Consequently, signal transduction marks a driving force in species formation, resulting in a wide range of signal transduction systems and their constituent signaling proteins across all branches of life.

In bacteria, signal transduction is dominated by so-called two-component signal transduction systems, consisting of a typically membrane-bound sensory histidine kinase and a cytosolic response regulator that, depending on its phosphorylation by the kinase, acts as a transcription factor for various downstream genes (Nixon et al., 1986; Stock et al., 2000). In distinction from this, other prokaryotic signaling pathways are sometimes referred to as one-component signal transduction systems (Ulrich et al., 2005), although they may likewise consist of a membrane-bound sensory component and a cytosolic effector. Rather than interacting directly, however, their components communicate by means of a cytosolic second messenger molecule that is produced by the sensor and recognized by the effector; in this, the release of a second messenger could be regarded equivalent to the phosphorylation of the response regulator.

In signal transduction systems with such layouts, the membrane-bound sensory component must fulfill three principal functions: (i) receive the original stimulus, (ii) convert it into a conformational change suitable to transduce information through the membrane, and (iii) create the intracellular signal for the cognate effector. Accordingly, many bacterial signaling proteins follow a modular architecture wherein each function is mediated by one or more specific domains that I will refer to along their putative role as input domains, transducer domains, and output domains. Depending on the signals produced by the output domain, signaling proteins are commonly classified into families, such as adenylate cyclases, guanylate cyclases, diguanylate cyclases, histidine kinases, and the antagonistic phosphodiesterases and phosphatases. Input and transducer domains, on the other hand, are frequently observed across family boundaries, e.g. the HAMP transducer domain, which was named for its presence in histidine kinases, adenylate cyclases, methylated chemotaxis proteins, and phosphatases. Thus, arguably, signaling proteins could equally be grouped according to their input or transducer domains, for example into HAMP domain-containing proteins or signaling proteins with a six-helical transmembrane domain.

In addition to the membrane-bound signaling proteins described so far, there exists a great number of cytosolic signaling proteins with only soluble input domains. Where known, these are activated by intracellular conditions, such as pH and light, or by crosstalk with other signal transduction pathways. This is well illustrated by CyaC from Arthrospira platensis, which likely originates from the fusion of an upstream histidine kinase to a downstream phosphorylation-activated adenylate cyclase, such that now the signal is intramolecularly propagated by an internal histidine kinase configuration phosphorylating the adjacent receiver (REC) domain (Kasahara and Ohmori, 1999). The bipartite nature of these signaling systems however suggests that soluble sensors evolved secondarily as an extension to the membrane-bound two-component signal transduction and second messenger systems; a view supported by the recent analysis of the evolutionary relationship

between membrane-bound and soluble members of the small molecule-binding domain (SMBD) family of input domains (Upadhyay et al., 2016).

Eukaryotes, on the other hand, differ considerably in the signals transduction systems they employ and also in the cellular responses that these systems control. Surprisingly, many of the widespread families of bacterial signal transduction proteins, such as diguanylate cyclases and histidine kinases, are absent or found in significantly smaller numbers, while others, such as serine/threonine kinases have significantly gained importance (Hanks and Hunter, 1995; Ponting et al., 1999). In animals, most extracellular conditions are sensed by the sizable family of G protein-coupled receptors (GPCRs), which communicate by the binding or release of G protein subunits, most importantly subunit α (Gsα, reviewed by Gilman, 1987). The various G protein subunits then trigger respective downstream signal transduction pathways. In the perhaps most important pathway thereof, Gsα stimulates the production of the intracellular second messenger cyclic adenosine monophosphate (cAMP), which in turn activates protein kinase A (PKA), a central determinant of eukaryotic cellular activity (Colledge and Scott, 1999; Acin-Perez et al., 2009; Turnham and Scott, 2016).

1.2 Adenylate cyclases produce the second messenger cAMP

cAMP is a general-purpose second messenger molecule mediating diverse cellular functions in most branches of life. In bacteria, cAMP signaling controls diverse metabolic processes, including the preference of a primary carbon source (catabolite repression), but mediates also various other roles, such as control of development, phototaxis, protein secretion, and virulence (Gersch et al., 1979; Süsstrunk et al., 1998; Stierl et al., 2011; McDonough and Rodriguez, 2011; Gancedo, 2013). Cellular cAMP acts mainly through binding and activation of the catabolite activator protein (CAP) (also referred to as cAMP receptor protein (CRP)), which functions as a transcriptional regulator for a wide range of genes (Figure 1A). In eukaryotes, on the other hand, cAMP acts mainly as a downstream second messenger

of GPCRs and is involved in many sensory and developmental processes (reviewed in Sadana and Dessauer, 2009). cAMP signals in eukaryotes are processed primarily by PKA, but also via several less-understood targets, such as cAMP-dependent calcium channels and cAMP-activated guanine nucleotide exchange factor (GEF) proteins (Figure 1B; Kaupp and Seifert, 2002; Bos, 2006). In addition, several non-canonical roles of cAMP have been reported for basal eukaryotes, as illustrated by the example of Dictyostelium discoideum secreting the molecule as an aggregation pheromone (Devreotes, 1989).

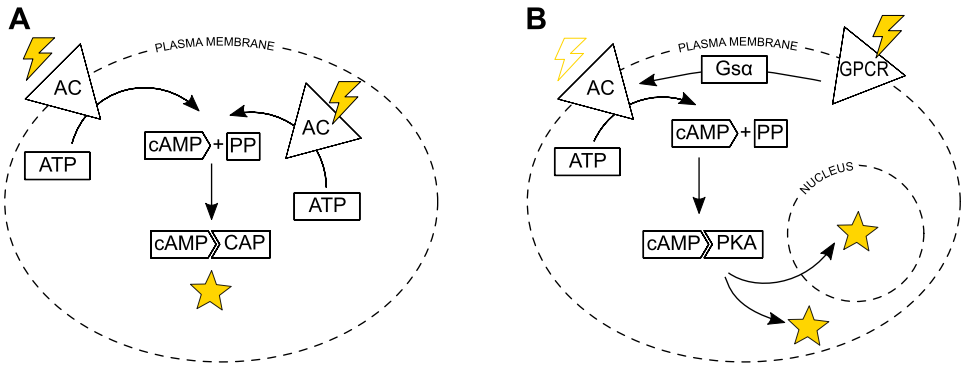


Figure 1: Schematic comparison of AC signaling in prokaryotes and eukaryotes. (A) Signal transduction in bacteria. cAMP production is thought to be regulated by direct stimulation of the ACs. Intracellular cAMP then activates the catabolite activator protein (CAP), which acts as a transcriptional regulator. (B) Signal transduction in eukaryotes. G protein-coupled receptors (GPCRs) function as primary sensors, which, by release of intracellular G protein subunit alpha (Gsα), activate downstream ACs. Intracellular cAMP levels are sensed primarily through protein kinase A (PKA), which initiates many downstream processes. A hypothetical receptor function of the eukaryotic membrane-bound AC (mAC), as derived from the comparison to prokaryotic systems, is indicated by a yellow outline. Note: for eukaryotes, signal transduction via the soluble ACs is left out for simplicity.

The cellular levels of cAMP depend on its production from ATP by adenylate cyclases (ACs) and its hydrolysis to AMP by phosphodiesterases (PDEs). The widespread abundance and physiological importance of cAMP could be taken as indication for a universal homology of cAMP-producing enzymes; surprisingly, however, this is not the case. Rather, there are at least six unrelated classes of ACs,

referred to by roman numerals as classes I-VI (Danchin, 1993; Sismeiro et al., 1998; Cotta et al., 1998; Téllez-Sosa et al., 2002). Five thereof are restricted to a narrow range of prokaryotic species and have not been studied in great detail. Class II stands out, in that these ACs do not act as signaling proteins but are secreted as toxins by a number of pathogenic bacteria, such as Bacillus anthracis and Bordetella pertussis (Hewlett and Wolff, 1976; Leppla, 1982). The existence of a putative seventh class is indicated by recent reports of AC activity in higher plants (Gehring and Turek, 2017). As of now, however, these proteins still await a more extensive characterization of their in vivo activity and substrate specificity.

Class III, finally, is the numerically largest, structurally and functionally most diverse, and pharmacologically most relevant; it is also the only one occurring in animals. Hence, in the following, the term AC specifically refers to class III, unless stated otherwise. The class III catalytic domain is related to the catalytic GGDEF domains of diguanylate cyclases and may have evolved from an early nucleotide polymerase (Figure 2; Artymiuk et al., 1997; Tesmer et al., 1997). The catalytic mechanism of these domains requires dimerization. Prokaryotic ACs form homodimers with two active centers on the subunit interface. Most eukaryotic ACs, in contrast, are so-called pseudoheterodimers, consisting of two complementary catalytic domains in a single protein chain that form only one active center between them.

Based on sequence similarity between the catalytic domains, class III has been further subdivided into four subclasses, termed IIIa to IIId (Linder and Schultz, 2003). Since the original definition, it has become clear that subclasses IIIc and IIId consist of several deeply-branching groups of equal rank, whose detailed classification needs revision (Dunin-Horkawicz and Lupas, 2010; Bassler et al., 2018). However, because there are only few relevant proteins in these groups, there is currently little incentive to do this. Here, I will address these groups collectively as subclass IIIc/d, which thereby forms the combined outgroup to subclasses IIIa and IIIb (see Figure 2, Bassler et al., 2018). In this revised classification, all three sub-

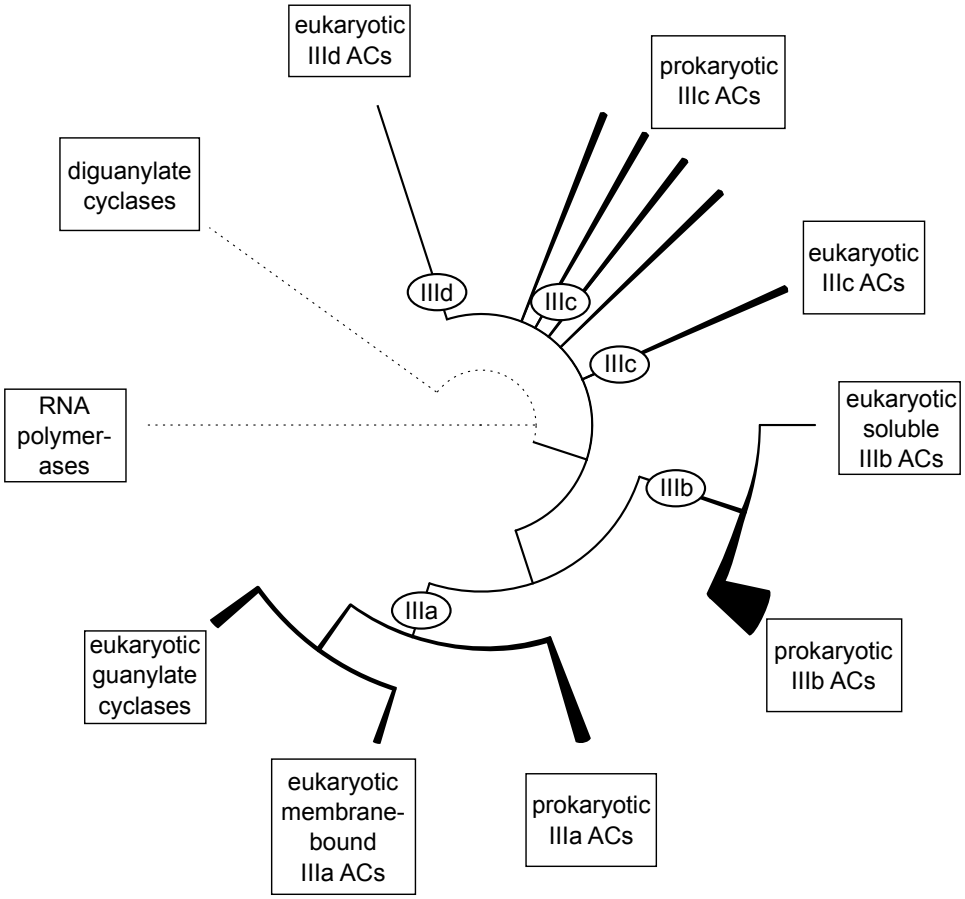


Figure 2: The evolution of class III ACs. Dashed lines represent distant homology to other protein families. Solid lines show relations between major subgroups of class III ACs. Line thickness within class III approximates the diversity in this branch. This figure has been reproduced from Bassler *et al.* (2018).

classes (IIIa, IIIb, and IIIc/d) contain both prokaryotic and eukaryotic proteins. Previously, subclass IIId had been the only subclass composed entirely of eukaryotic members. The medically relevant ACs of animals, however, are found entirely in subclasses IIIa and IIIb, while the eukaryotic members of subclass IIIc/d mostly come from protists and fungi.

In addition to canonical ACs, class III contains a significant number of proteins that appear to have shifted their substrate specificity from ATP to GTP, resulting in class III guanylate cyclases (GCs). By far the most class III GCs are found in multi-

cellular eukaryotes, where they constitute a significant family of signaling proteins in its own right. Reports of GCs from protozoan organisms (e.g. Linder et al., 1999; Roelofs et al., 2001; Baker, 2004) and possibly also from bacteria (Linder, 2010; Marden et al., 2011; Ryu et al., 2015) indicate that GTP substrate specificity might have evolved independently on multiple occasions (Roelofs et al., 2001; Bassler et al., 2018). While the downstream cellular processing and physiological roles of cAMP and cGMP differ greatly, a comparison between class III ACs and GCs may yield helpful insights into the possibly conserved mechanism of intramolecular signal transduction in both families. Thus, when I refer to class III GCs in this work, it is for the larger context of all class III nucleotide cyclases without explicit considerations of the physiological differences of cGMP signaling.

1.3 Bacterial ACs are modular multidomain proteins

Bacterial ACs are structurally highly diverse multi-domain proteins, in which they arguably reflect the complexity of signals conveyed by them. Much of this diversity seems to originate from the modular recombination of input and transducer domains in adaptation to particular signaling requirements, as is evident for example by the frequent occurence of members of the SMBD family in many bacterial ACs (Anantharaman et al., 2001; Shenoy and Visweswariah, 2004; Schultz and Natarajan, 2013). At the same time, only few ACs have been characterized experimentally, and where they have been studied, it was often only the catalytic domains and under primarily biochemical considerations (e.g. Yan et al., 1997; Zhang et al., 1997; Tesmer and Sprang, 1998; Tesmer et al., 1999; Sinha et al., 2005). The factors regulating the activity of full-length ACs under physiological conditions, on the other hand, have so far been identified for only three, all cytosolic, proteins: CyaB1 of Anabaena sp., which is activated upon binding of cAMP to the N-terminal tandem GAF domains (Figure 3; Kanacher et al., 2002); the mycobacterial Rv1264, which is activated by acidic conditions sensed in an N-terminal pH sensor domain (Figure 3; Linder et al., 2002; Tews et al., 2005); and two instances of photosen-

sitive ACs from Oscillatoria and Begiatoa, termed oaPAC and bPAC, respectively, which are activated by an FAD cofactor bound to their N-terminal BLUF domains (Figure 3; Ohki et al., 2016; Lindner et al., 2017).

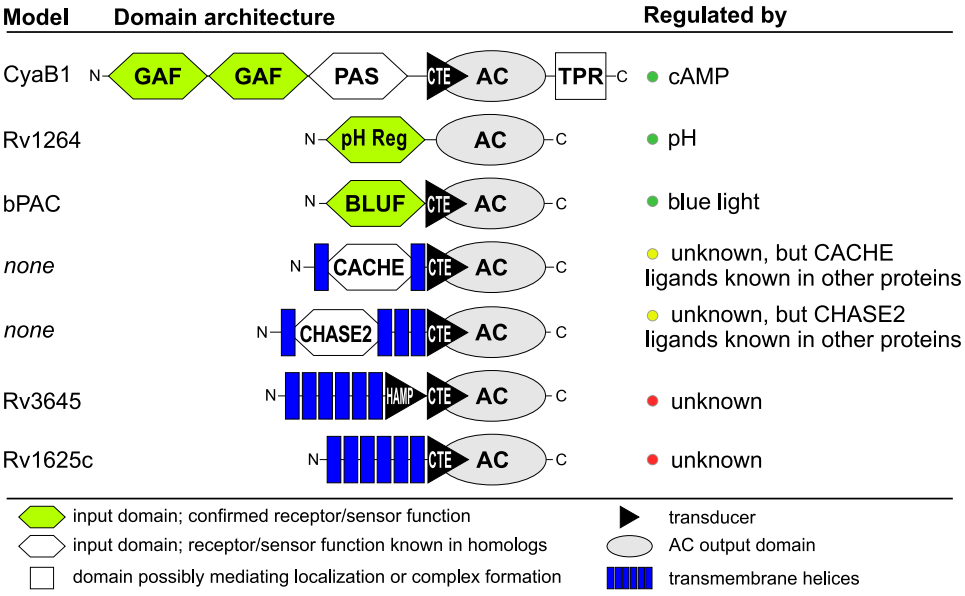


Figure 3: Relevant AC domain architectures. AC domain architectures as discussed in the text. Where ACs of this type have been studied, the name of the best-studied model protein is provided. Right column indicates whether regulatory conditions have been identified for this protein (green dot), for homologs of the putative input domains in other signaling proteins (yellow dot), or not at all (red dot). For known regulatory conditions, the sensing input domains are highlighted in green. Note that this domain annotation includes the cyclase transducer element (CTE), which was only discovered as part of this work.

A similar mode of activation may be expected for several bacterial ACs, whose N-terminal domains have not been characterized in ACs, such as PAS domains or single GAF domains, but for which a sensory function has been established in other families of signal transduction proteins (Anantharaman et al., 2001; Hefti et al., 2004; Karniol et al., 2005; Cann, 2007; Henry and Crosson, 2011; Fushimi et al., 2017). The same applies to two widespread types of membrane-bound ACs characterized by transmembrane domains with two and four membrane-spanning helices, whose sizable extrcellular loops are typically recognized as CACHE and CHASE2

domains, respectively (Figure 3; Anantharaman and Aravind, 2000; Mougel and Zhulin, 2001; Zhulin et al., 2003). Interestingly, for a third widespread category of bacterial membrane-bound ACs, which is defined by six membrane-spanning helices (6TM), a sensory function had been proposed based on positional considerations, but had neither been documented within ACs nor in homologs of other signaling proteins (Nikolskaya et al., 2003). In the investigation of this function, the 6TM AC Rv1625c from Mycobacterium tuberculosis has played an important role as a model protein for both bacterial and eukaryotic ACs (see section 1.6).

Several studies have sought to classify the structural diversity ACs, not least in the interest of a universal ontology that could clarify the role of an individual protein within its cellular signaling network (e.g. Linder and Schultz, 2003; Shenoy and Visweswariah, 2004; Schultz and Natarajan, 2013; Bassler et al., 2018). A common principle between many ACs appears to be that putative input domains are located at the N-terminus, while the catalytic output domain is found at or near the C-terminus. Notably, this prototypical arrangement is also found in several other families of bacterial signaling proteins, including histidine kinases. An important, yet unsolved question in this regard is how the observed multiplicity of sensory domains is able to communicate with the conserved catalytic domain. The structures of two ACs have so far been determined as a whole, enabling the study of the interactions between their input and catalytic domains. The pH-regulated Rv1264 has been solved in active and inactive conformations, revealing the linker between the pH sensor domain and the catalytic domain as the site of major structural rearrangements between both forms (Tews et al., 2005, see Figure 4). In the inactive conformation at neutral pH, this linker forms a long α-helix, hindering the dimerization of the catalytic domains. Under acidic conditions, this helix collapses, allowing the catalytic domains to form a productive homodimer.

For bPAC, the second AC with known full-length structure, only one conformation is available. However, hydrogen-deuterium exchange data likewise indicates the helical linker between the sensory domain and the catalytic domain as the site

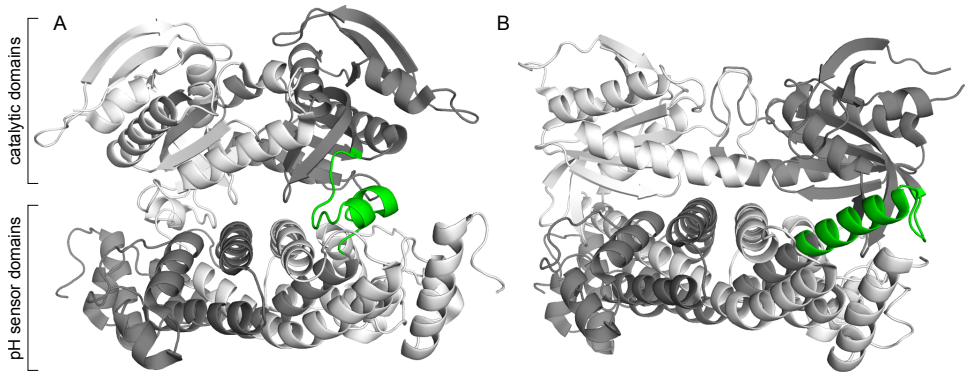


Figure 4: Structures of the mycobacterial Rv1264 in active and inactive conformation. (A) Rv1264 in its active conformation at acidic pH (PDB: 1y11); (B) inactive conformation at neutral pH (PDB: 1y10). The linker region between the regulatory pH sensor domain and the catalytic domain in one chain is highlighted in green. The two chains were colored in gray and white for better distinction.

of greatest structural change between the active and inactive state (Lindner et al., 2017). Since in bPAC the linkers of both chains interact in a two-helical coiled coil, as opposed to the solvent-exposed single α-helices of Rv1264, the activity-determining conformational change likely differs. We have found that the linkers of Rv1264 and bPAC represent two fundamental types of AC structure with profound influence on the regulation of these proteins, and that the presence of the coiled-coil linker could serve as indicator for regulatory signals being passed from the N-terminus towards the catalytic domain (see Results and Discussion).

1.4 Eukaryotic mACs resemble bacterial 6TM ACs

Compared to their bacterial counterparts, eukaryotic ACs show considerably lower diversity and their domain architectures are frequently conserved throughout major clades; for example, in all animals, only two architectural types are present: a membrane-bound AC (mAC) of subclass IIIa and a soluble AC (sAC) of subclass IIIb (Figure 5). Particularly the mACs are of high medical relevance because of their role as downstream messengers of G protein-coupled receptors, which connects them to a great number of sensory and developmental processes (Dessauer

et al., 2017). Like most eukaryotic ACs, the animal mACs are pseudoheterodimers, consisting of two complementary subunits in a single chain. Each subunit is composed of a six-helical transmembrane (6TM) domain and a subclass IIIa catalytic domain. In this, they closely resemble the domain architecture of several bacterial 6TM ACs; e.g. the mycobacterial Rv1625c, which has been proposed as a bacterial model system for mammalian mACs (Guo et al., 2001).

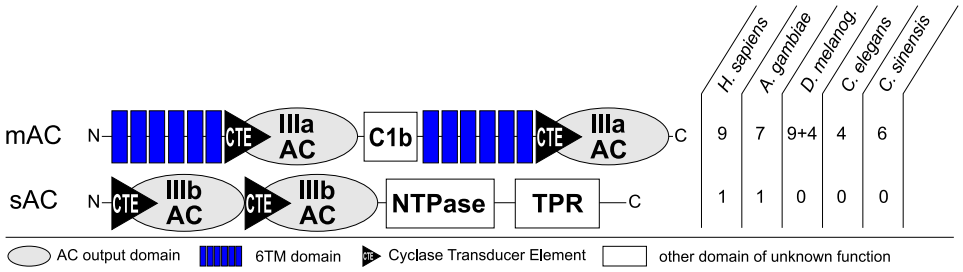


Figure 5: Domain architectures of animal ACs. All animals possess only two architectural types, termed membrane-bound ACs (mACs) and soluble ACs (sACs). In the membrane-bound form, the conserved linker between the subunits of the pseudoheterodimer is sometimes considered a domain and referred to as C1b. The soluble form carries two C-terminal domains with sequence similarity to NTPase domains and the TPR motif, respectively, for which the function is largely unclear. Numbers on the right indicate genomic copies of this AC type in the for *Homo sapiens*, *Anopheles gambiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, and *Clonorchis sinensis*. Note that this domain annotation includes the CTE, which was only discovered as part of this work.

The mAC type appears to have undergone extensive lineage-specific expansion, as is evident from the varying genomic copy numbers between animal genomes (Figure 5; Bassler et al., 2018). Mammals, including humans, possess nine mAC paralogs, commonly referred to as AC1-AC9, which differ in their expression patterns and, surprisingly, in their regulation through cytosolic factors. All mACs are stimulated by Gsα and, except for AC9, by the plant diterpene forskolin. In addition, several factors have been reported that act only on specific mAC paralogs, for example the G protein subunits Gαi and G$\beta\gamma$, Ca^{2+} and calmodulin, and protein kinases A and C (comprehensively reviewed in Dessauer et al., 2017). Where known,

these factors are bound in the intracellular parts of the proteins (Tesmer et al., 1997; Wu et al., 1993; Vorherr et al., 1993; Chen et al., 1995; Weng et al., 1996). A function of the sizable 6TM domains beyond cellular localization, on the other hand, has remained elusive.

1.5 AC 6TM domains could be regulatory receptors

Already in the initial description of an AC peptide sequence (the bovine mAC1), the authors expressed their surprise about the presence of two sizable 6TM domains, which they deemed unexpectedly complex given the known intracellular regulation by Gsα (Krupinski et al., 1989). Since then, several hypotheses regarding the function of the AC 6TM domains have been put forward, variously proposing a role as solute transporter or ion channel (Krupinski et al., 1989), as voltage sensors (Reddy et al., 1995), or as dimerization factors (Cooper and Crossthwaite, 2006). None, however, has found experimental validation. Instead, mammalian mAC research has focused on their regulation by cytosolic factors, most prominently via GPCRs, while the question of a role of the 6TM domains beyond mere cellular localization remained inconclusive.

6TM domains are also present in several bacterial ACs, for example in Rv1625c and Rv3645 from M. tuberculosis. Intriguingly, the position of the 6TM domains in these proteins closely resembles that of two-helical and four-helical transmembrane domains, for which a receptor function by means of their lengthy extracellular loops is widely accepted (see Figure 3; Anantharaman and Aravind, 2000; Zhulin et al., 2003). Further, in Rv3645, the 6TM anchor is connected to the catalytic domain by a HAMP transducer, which was shown to propagate regulatory information in constructs with a heterologous sensor domain (e.g. Hulko et al., 2006; Kanchan et al., 2010; Ferris et al., 2011; Mondéjar et al., 2012). While this suggested that at least some 6TM domains could be sites of regulatory input, an actual receptor function was largely rejected due to the absence of solvent exposed loops of significant length, which were considered necessary for the coordination of sufficiently

complex ligands.

This assumption of necessity was however challenged by the recent identification of the hydrophobic ligands of several bacterial quorum sensor histidine kinases (Higgins et al., 2007; Spirig et al., 2008). These quorum sensors possess an N-terminal 6TM domain with a topology similar to AC 6TM domains, particularly with respect to the absence of significant extracellular loops. It is strongly believed that the ligand is recognized by the 6TM domains inside the membrane due to the hydrophobicity of the ligand and mutation analyses of the anchors (Ng and Bassler, 2009; Tiaden and Hilbi, 2012). Indicating that the 6TM domains of ACs could possibly coordinate small molecules in a similar fashion, this raised the hypothesis that some or even all AC 6TM domains could function as regulatory receptors. However, with the type of putative ligand completely unknown, its effect (stimulation or inhibition) unclear, and comparably little established methodology for working with AC transmembrane domains or full-length proteins, a direct approach to the problem seemed impossible.

1.6 Chimeras between quorum sensors and ACs are regulated by the quorum sensor 6TM receptor domain

In order to test whether bacterial 6TM ACs are fundamentally compatible with regulatory information from an integral transmembrane receptor, Joachim Schultz and coworkers at the University of Tübingen constructed a chimera between the quorum sensor CqsS from Vibrio harveyi and the mycobacterial AC Rv1625c. Several such chimeric constructs between signaling proteins have been constructed before, e.g. between the Escherichia coli aspartate receptor Tar and the osmolarity receptor EnvZ (Utsumi et al., 1989), between the nitrate-sensing histidine kinases CpxA and NarX (Appleman et al., 2003), and between Tar and the mycobacterial 6TM AC Rv3645 (Hulko et al., 2006). These chimeras differed from the CqsS-Rv1625c one in two critical aspects: (i) the input domain of previous chimeras was always a two-helical transmembrane domain that sensed the respective ligand through an

extracellular loop, and (ii) in all constructs, the parent proteins contained a HAMP transducer domain, which marked a suitable point of recombination. The CqsS-Rv1625c chimera, in contrast, used the 6TM receptor domain from CqsS, which was expected to bind the ligand inside the membrane, and no homologous element between the parent proteins indicated a common interface that could facilitate signal transduction in the recombinant protein.

Surprisingly, the chimera between CqsS and Rv1625c was functional and its AC activity could be regulated using the natural ligand of CqsS, CAI-1 (Beltz et al., 2016). A subsequent analysis of linkage points in a similar chimera between the quorum sensor LqsS from Legionella pneumophila and Rv1625c revealed that successful coupling was only possible in one specific position, while all other attempts led to chimeras that were either inactive, unregulated, or both (Ziegler et al., 2017). This indicated that the interaction between the quorum sensor domain and the AC catalytic domain was specific and that Rv1625c was able to decode the intramolecular signal coming from the 6TM receptor. Naturally, this spurred the hypothesis that the natural 6TM domain of Rv1625c could likewise be a transmembrane receptor. It also raised the question to what extent this observation represented a general principle that could be transferred to other 6TM domains and ACs, and particularly to the 2x6TM domains of the pharmacologically highly significant mammalian AC pseudoheterodimers.

In this thesis, I will try to answer these questions, complementing the exciting experimental data by the Schultz group with a bioinformatic analysis that seeks to deduce the underlying principles from observations made in the Rv1625c model protein.

2 | Objectives

Six-helical transmembrane (6TM) domains feature in many prokaryotic and eukaryotic adenylate cyclases (ACs). The widespread presence of this domain type is increasingly understood as indication that it serves a more complex function than mere cellular localization, for example as an input domain to regulatory stimuli. However, direct evidence for this function is lacking at present. A first step towards the experimental investigation of the hypothesized function was made with the successful regulation of AC activity by means of a heterologous 6TM receptor domain (Beltz et al., 2016; Ziegler et al., 2017), but effectively this only established the general possibility of the concept.

Proving the actual existence of this function in 6TM domains will however require the identification of natural ligands, which seems extremely challenging at present given the overall lack of knowledge on AC 6TM domains. Notably, this includes the absence of any information on their structures, the location of their putative ligand binding sites, the mechanism of intramolecular signal transduction, and the regulatory effect (i.e. stimulation or inhibition), rendering any informed approach towards ligand identification impossible. At the same time, no suitable assay to scan for the binding of ligands within the membrane is available, ruling out a high-throughput approach as well.

In this context, we have joined the effort, complementing the experimental work in the Schultz group with extensive sequence analyses that seek to generalize the findings in the Rv1625c model system and provide reference points for new experimental approaches. Concretely, this meant to work towards two major aims:

(i) A classification of AC 6TM domains. At the beginning of this work, the pairwise relationship between AC 6TM domains was completely unclear. The 6TM designation had been chosen on entirely topological grounds and was arguably kept

alive by the intricacies of remote homology detection between sequences of integral transmembrane domains. In order to clarify the extent by which functional insights could be transfered between 6TM domains, and also to identify conserved features from which functional predictions could be made, it was necessary to establish a homology-based classification of these domains.

(ii) Investigation of the intramolecular signal transduction between upstream regulatory domains and the AC catalytic domain. The analysis of linkage points between the LqsS receptor domain and Rv1625c had revealed the importance of a short segment on the N-terminus of the catalytic domain for the regulation of AC activity. We aimed to understand the precise role of this linker segment and how it enabled the regulation of AC activity by means of a heterologous 6TM receptor.

3 | Results and Discussion

3.1 Classification of AC 6TM domains

The results summarized in this section were published in Beltz et al. (2016). The full manuscript and a detailed list of my contributions can be found in appendix A.

Membrane-bound adenylate cyclases (ACs) are frequently classified according to their number of membrane-traversing helices, grouping them into ACs with two-helical, four-helical, and six-helical transmembrane (6TM) domains. Unlike two-helical and four-helical transmembrane domains, where the annotation of extracellular loop regions offered a suitable criterion for a more detailed, homology-based subclassification, the pairwise relationship between AC 6TM domains had not been explored systematically before this work. With recent experimental results indicating that domains of this layout could function as regulatory receptors, a more detailed classification of 6TM domains was necessary.

Six conserved types of AC 6TM domains

In order to provide a comprehensive overview of AC 6TM domains, I extracted approximately 1,600 6TM domains from more than 5,500 predicted AC sequences in the taxonomically balanced Uniprot reference proteomes database (release 04/2015; the UniProt Consortium, 2012). The resulting 6TM domain sequences were subjected to a cluster analysis using highly sensitive pairwise profile hidden Markov model comparisons (Frickey and Lupas, 2004; Remmert et al., 2012). This identified six major groups of AC 6TM domains (Figure 6). All groups had characteristic sequence profiles, indicating they represented conserved 6TM domains or domain

subtypes.

Two groups originated mostly from eukaryotic ACs and corresponded to the 6TM domains in the N-terminal and C-terminal subunits of the mAC pseudoheterodimers (Figure 6A and B). These clusters contained the sequences from all mAC paralogs, highlighting the homology between all eukaryotic mAC 6TM domains. Interestingly, the cluster otherwise associated with the N-terminal subunit of eukaryotic mACs also contained a small number of 6TM domains from multicellular cyanobacteria, such as Nostoc or Anabaena; the implications of this observation for the evolution of mAC 6TM domains are not clear.

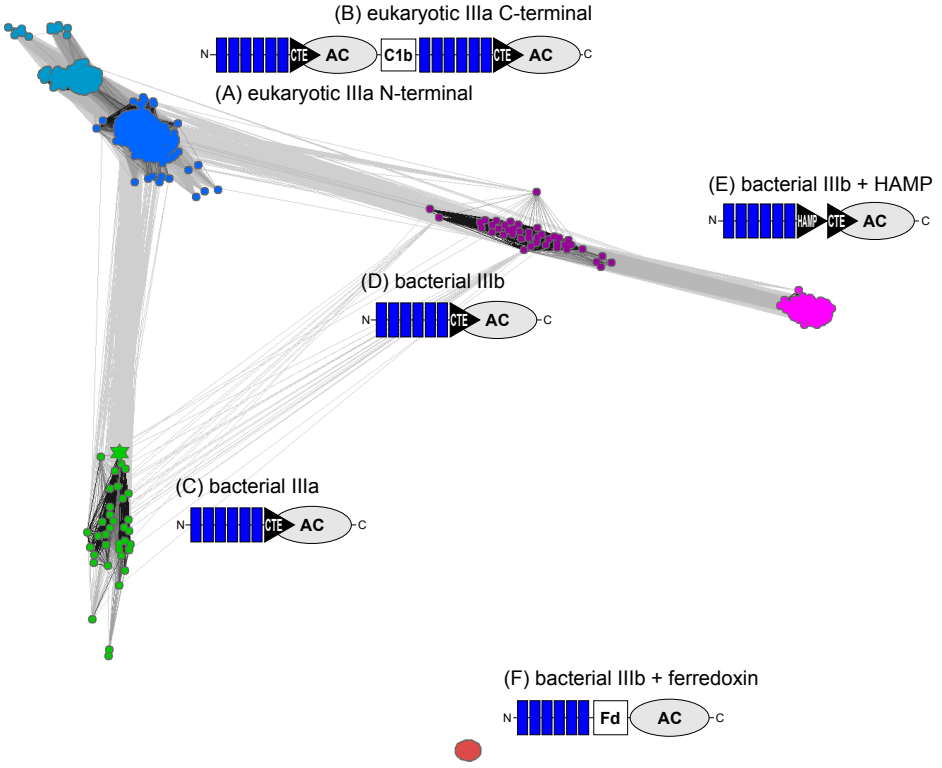


Figure 6: Cluster analysis of 6TM domains in ACs. Each dot represents a single 6TM domain. HHsearch hits better than p=0.0005 are shown as connecting lines; darker line color indicates a better hit. The position of Rv1625c is highlighted by a green star. Pictograms next to clusters indicate the domain architectures of proteins with this 6TM domain (compare Figure 3). Note that the domain annotations include the cyclase transducer element (CTE), which was not known when the original manuscript was published.

The remaining four groups consisted entirely of bacterial ACs, but were not separated along any recognizable taxonomic division. Rather, the grouping correlated with properties of the full proteins, particularly with their AC subclass and domain architecture. The groups corresponded to: one group of subclass IIIa, in which the proteins consisted of only the 6TM domain and the catalytic domain (Figure 6C), and three groups of subclass IIIb, which were characterized by a domain architecture consisting only of the 6TM domain and the catalytic domain (Figure 6D), by the presence of a HAMP domain (Figure 6E), and by the presence of a peculiar ferredoxin domain (Figure 6F), respectively. The mycobacterial Rv1625c, which had been used for the chimeric constructs with bacterial quorum sensors, located in the cluster of bacterial IIIa ACs (highlighted by a green star in Figure 6), while the previously studied Rv3645 clustered with the group of HAMP domain-containing ACs of subclass IIIb. We understood this correlation between features of the membrane anchors and the full-length protein as evidence for a co-dependence between the 6TM domains and other parts of the protein. As such, this observation clearly supports the notion that the 6TM domains possess a more complex function beyond cellular localization. Further, the presence of a HAMP domain strongly indicates the transduction of regulatory information from the N-terminal 6TM domains towards the catalytic domain in at least one group (Dunin-Horkawicz and Lupas, 2010).

The relationship between the different groups of AC 6TM domains was not fully clear. All groups except for the one characterized by a ferredoxin domain were connected by relatively weak, yet consistent HHsearch hits (indicated by faint lines in Figure 6). These hits were of low statistical significance (most p-values > 10^{-5}) and could not be attributed to a clearly identifiable sequence pattern conserved between all five groups. Further, we noted that the supposedly unrelated 6TM domains of bacterial quorum sensors could be aligned to these groups at similar confidence, but not to the outlying ferredoxin ACs (see Figure A7). At the time, we were unsure how to interpret these results, mostly because only very few reference cases

for remote homology between integral transmembrane domains are known. If understood as an indication for true sequence similarity, the hits may have suggested remote homology between several types of AC 6TM domains and also the 6TM receptors of bacterial quorum sensors. This may have implied a conserved function between these domains, particularly the possibility to recognize hydrophobic ligands inside the membrane. On the other hand, the hits may have been artifacts, resulting from the similar topology of these domains and the reduced amino acid alphabet of transmembrane sequences. Meanwhile, with additional insights into the evolution of AC domain architectures and more experience with the comparison of transmembrane sequences, I think that homology between these domains is becoming increasingly plausible (see Conclusion and Outlook).

Surprisingly, the minimum-sized loops, which we had considered a universal hallmark of AC 6TM domains, were only present in three clusters (the ones originating from bacterial IIIa ACs, from bacterial subclass IIIb ACs without a HAMP domain, and from the N-terminal eukaryotic subunits). All other groups contained at least one conserved, significantly solvent-exposed loop (see appendix Figures A.7 and A.7-S1). The function of these loops is unknown.

The 6TM domains in the ferredoxin-containing cluster stood out particularly. In this group, the lengths of the transmembrane spans differed significantly from all other groups, suggesting different membrane-crossing angles of the helices, and hence a different fold. This cluster was also the only one with a homolog of known structure, being clearly related to the heme-binding cytochrome B subunits of fumarate reductases/succinate dehydrogenases. The fumarate reductases/succinate dehydrogenases coordinate heme through four histidine residues in their transmembrane domains. Notably, these histidine residues are conserved in the AC 6TM domains. Together with the predicted ferredoxin domain between the 6TM domain and the catalytic domain, this suggests an electron transport-related function of these ACs, for example as sensors of the redox state. This hypothesis is currently under investigation in the groups of Prof. Gottfried Unden at the University

of Mainz and Prof. Joachim Schultz at the University of Tübingen.

3.2 The cyclase transducer element

The results summarized in this section were published in Ziegler et al. (2017). The full manuscript and a detailed list of my contributions can be found in appendix B.

In the chimera between the bacterial quorum sensor CqsS and Rv1625c (see Introduction, section 1.6), the AC was connected at arginine 218. This position was chosen for its homology to the connection point of previous chimeric constructs (Winkler et al., 2012; Beltz et al., 2016). A subsequent, systematic inquiry of linkage points in chimeras between the Legionella quorum sensor LqsS and Rv1625c revealed that this position was of critical importance for the functional interaction of both parts (Ziegler et al., 2017). When the connection was made upstream of arginine 218, the resulting chimera had comparatively high basal activity but lost all sensitivity to the LqsS ligand. Chimeras connected up to four residues downstream of this position, on the other hand, maintained a mild stimulatory response but progressively lost activity (Ziegler et al., 2017). In comparing this position between homologs, we realized that arginine 218 marked the beginning of a short segment with elevated sequence conservation. The segment reached out for 19 residues downstream, terminating in a clear drop of sequence conservation and secondary structure propensity (Figure 7). Due to its conservation and importance for the propagation of intramolecular signals, we named this segment the cyclase transducer element (CTE) and proposed that it could play a central role in the integration of regulatory signals.

CTEs are found in ACs of subclass IIIa and IIIb

In order to assess the abundance of this element across different ACs, I comprehensively analyzed the N-termini of AC catalytic domains for CTE-like sequences.

Surprisingly, this found that CTEs were universally present in ACs of subclass IIIa, including the animal mACs and also the eukaryotic guanylate cyclases (appendix Figure B.3). A cluster analysis of vertebrate CTE sequences further showed minor, yet characteristic sequence divergence along their origin from a particular mAC paralog and subunit of the pseudoheterodimer (appendix Figure B.5). This indicated an adaptation and continued co-evolution between the CTE and its adjacent 6TM and catalytic domain. Notably, the CTE thus represented the first clearly homology-based indication for a similar regulation of bacterial and eukaryotic ACs through upstream domains.

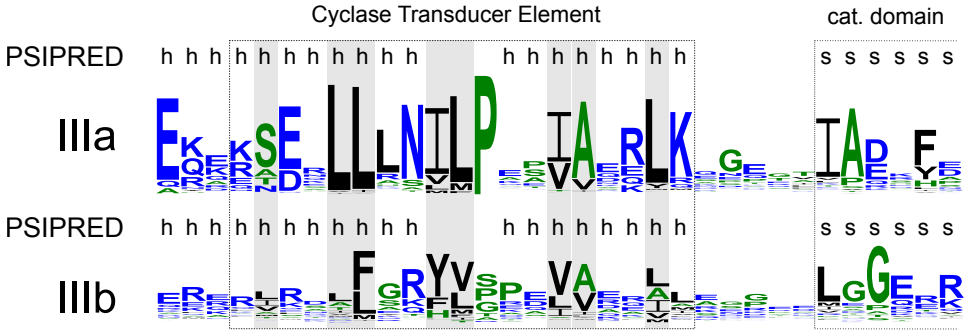


Figure 7: Sequence logos of CTEs in AC subclasses IIIa and IIIb. The PSIPRED lines indicate the secondary structure as predicted by PSIPRED (Jones, 1999) - h: α-helix, s: β-strand. Conserved hydrophobic positions in the CTE that could indicate coiled-coil core residues are highlighted in gray. This figure is a simplified replotting of Ziegler *et al.* (2017) Figure 3.

CTEs were also identified in ACs of subclass IIIb, including the soluble ACs (sACs) of animals. However, several other ACs in this subclass clearly lacked a CTE, for example the ferredoxin-containing ACs described earlier (Beltz et al., 2016; see section 3.1). It was also absent from all ACs of subclass IIIc/d, such as mycobacterial Rv1264, where another type of linker appears to mediate the regulation of catalytic activity (Tews et al., 2005). I will discuss a structural reason for the absence of CTEs in these proteins in the context of my global analysis of AC domain architectures (section 3.3).

The CTE controls the dimerization of the catalytic domains

The CTE was predicted to be of helical secondary structure. In subclass IIIa, the helix was locally disturbed by a central, almost invariant proline residue. In subclass IIIb, this proline was largely absent; surprisingly, however, secondary structure prediction nonetheless indicated a local disruption of the helix (Figure 7). Surrounding this position, the CTE showed a conserved pattern of hydrophobic residues in every third or fourth position, indicated coiled-coil interactions between the CTEs of the AC dimer (Figure 7). The exact coiled-coil register was, however, ambiguous and could not be assigned confidently.

Searching for CTE homologs in proteins of known structure, we found two cases of class III nucleotide cyclases, in which the CTE was included in the solved fragment although its functional relevance had not been recognized at the time. One structure originated from the soluble β1 guanylate cyclase from rat, where the CTE had been solved together with the upstream coiled-coil signaling helix (Ma et al., 2010). The other was from the human sAC, where the CTE had been solved as part of the upstream catalytic domains (Kleinboelting et al., 2014; Figure 8A). Both structures showed the element folded into short helices on both ends, separated by a central bend of approximately 45°. However, only the N-terminal helix made the expected coiled-coil contacts to its respective symmetry partner in the dimer, whereas the C-terminal helices extended into opposite directions before transitioning into the globular catalytic domain (Figure 8A).

Given the general ambiguity of the coiled-coil register and the surprisingly irregular packing of hydrophobic side chains in the C-terminal helix (Figure 8B), we conjectured that CTE may switch between multiple conformational states. Similar cases of sequence ambiguity and conformational flexibility have been observed before in other coiled-coil elements of signal transduction proteins, for example in the two-helical signaling helix (Anantharaman et al., 2006) and the four-helical HAMP domain (Hulko et al., 2006). Further, we were guided in our considerations by the parallels between the CTE and the dimerization and histidine phosphotrans-

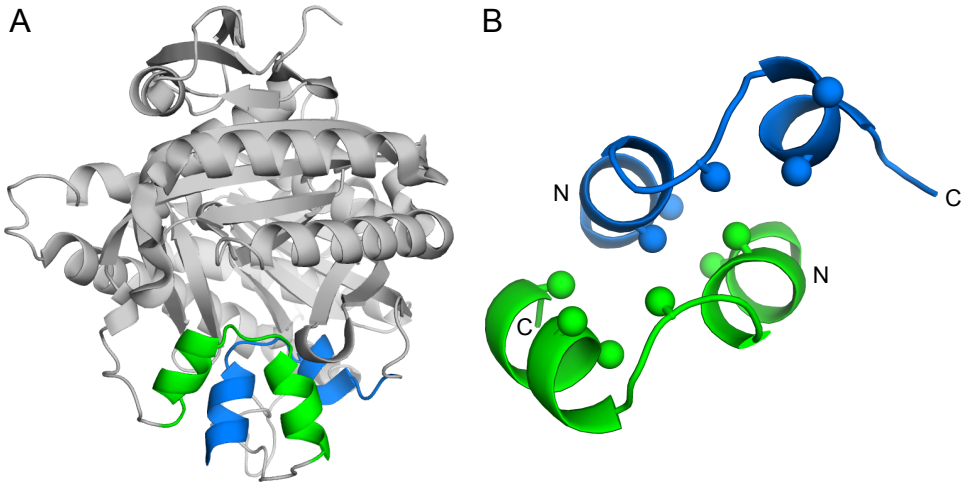


Figure 8: CTE crystal structures. (A) Structure of the catalytic domains and CTEs in the human soluble AC (PDB: 4clf). The CTEs are highlighted by green and blue color. (B) Top view onto the CTE dimer in the human soluble AC. The Cβ atoms of conserved hydrophobic residues are drawn as spheres.

fer (DHp) domain of histidine kinases. Both are coupled on their N-terminus to various input or transducer domains, such as the aforementioned signaling helix or HAMP domain. And, depending on signals from N-terminal domains, both set the activity of their respective output domains.

Based on this analogy, we illustrated a possible mechanism for the conformational changes in the CTE during signal transduction. In many ACs, the CTE directly extends the coiled coil of an upstream signaling helix or HAMP domain, and is therefore likely to undergo the same conformational change. We envisaged that this movement of the N-terminal helix triggers a repacking of the C-terminal helix, such that it forms the expected coiled-coil interactions with its counterpart in the opposite CTE (see Figure 9). Thereby the CTE could stretch out and adopt a straight conformation. By extrapolation from the linker region in the mycobacterial Rv1264 (Tews et al., 2005, see Figure 4), the more constrained conformation of these would correspond to the inactive state, while the less constrained conformation allows the catalytic domains to dimerize and thus represents the active state. Between these extremes, we expected one or several asymmetric states in which

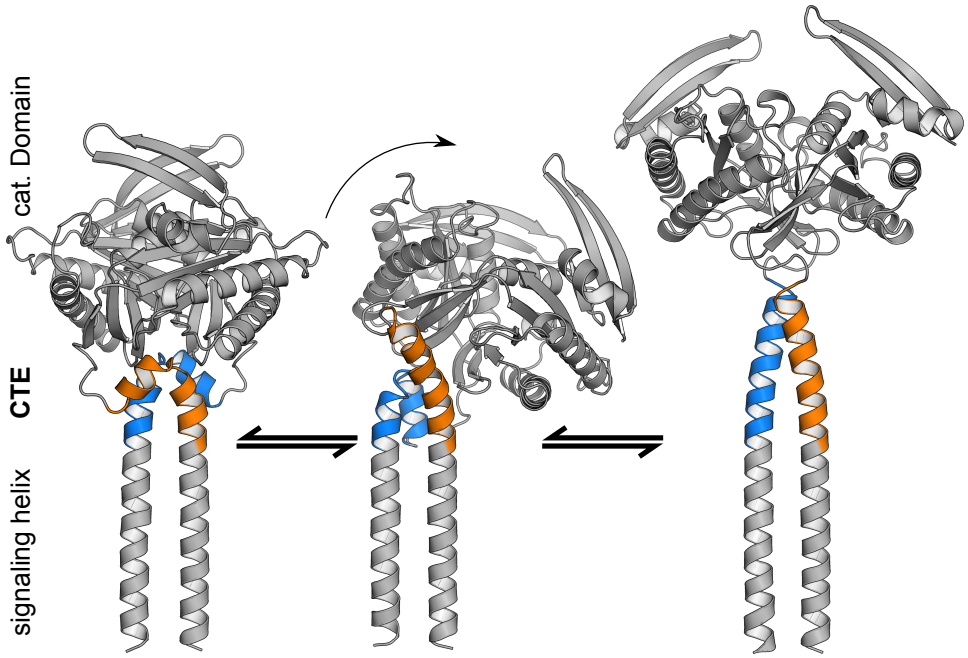


Figure 9: Computational models of the CTE in different states. The CTE was modeled together with the C-terminal catalytic domain and an N-terminal signaling helix (see Ziegler *et al.*, 2017). Left: Both CTEs are in the experimentally observed bent conformation. Center: The transition progresses through an asymmetric state in which one CTE adopts the hypothetical straight conformation while the other remains bent. Right: Both CTEs adopt the hypothetical straight conformation by interaction of the hydrophobic residues in the N-terminal helix. This figure has been adapted from Ziegler *et al.* (2017).

one CTE adopts the straight conformation and the other remains bent. Comparable asymmetric states were found in the DHp domain and are suggested by the sequence divergence between the N-terminal and the C-terminal CTE in eukaryotic pseudoheterodimers. Taken together with the conserved differences between the N-terminal and C-terminal 6TM domains of eukaryotic mAC, this could indicate that the (pseudo-) heterodimerization of these proteins enabled them to decode more complex, possibly asymmetric signals.

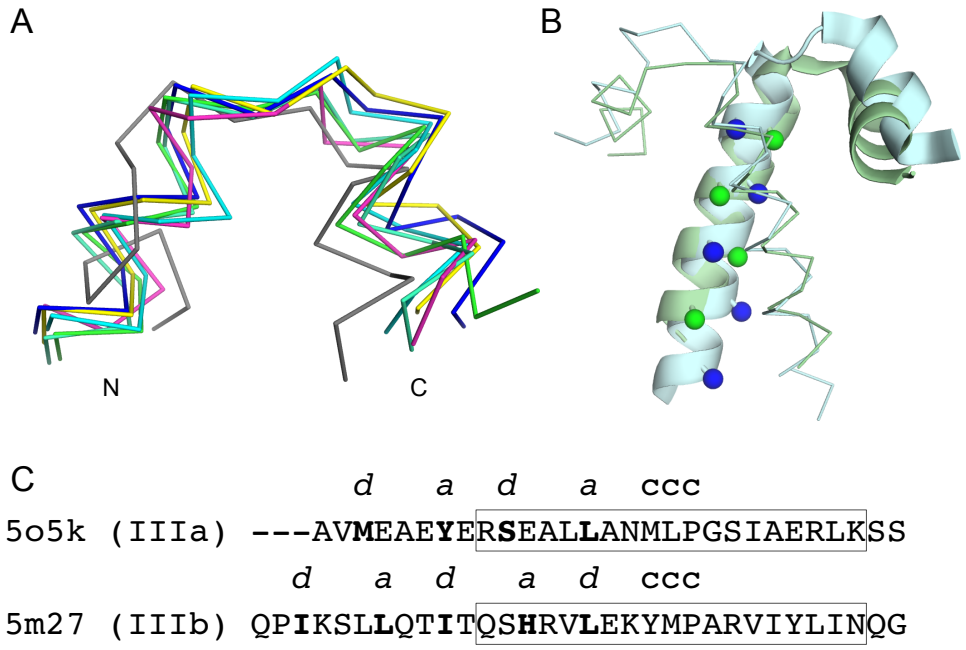


Figure 10: Comparison of CTE structures and sequence registers. (A) Superposition of all currently available CTE structures: the N-terminal and C-terminal CTEs of the human soluble AC (green and lime; PDB: 4clf), the CTE of a β1 guanylate cyclase from rat (cyan, PDB: 3hls), the CTE of a bacterial subclass IIIa AC from *Mycobacterium intracellulare* (magenta, PDB: 5o5k), the CTEs of bacterial subclass IIIb AC from *Beggiatoa* sp. (yellow, PDB: 5m27) and *Oscillatoria acuminata* (blue, PDB: 4yus), and a probably misfolded CTE of a subclass IIIb AC from *Synechocystis* (gray, PDB: 6fht). (B, C) Different coiled-coil registers in the CTEs of bacterial ACs from subclass IIIa (green, PDB: 5o5k) and IIIb (blue, PDB: 5m27). The Cβ atoms of coiled-coil core residues are shown as spheres. Labels *a* and *d* denote the heptad position, ccc marks the position of the central bent.

Recent developments

Since our discovery of the CTE, several other groups have also published results that support a central role of the CTE in the regulation of AC activity. Notably, this includes four more CTE structures from bacterial ACs of subclass IIIa and IIIb (Ohki et al., 2016; Lindner et al., 2017; Vercellino et al., 2017; Etzl et al., 2018). All except for one structure showed the CTE in the already known conformation with two helices separated by a central bend (Figure 10A). The single deviant CTE

was probably connected in the wrong coiled-coil register and the activity of its catalytic domain was very low (Etzl et al., 2018). Similar delirious effects were also noted when several disease-associated mutations found in the CTEs of human nucleotide cyclases were introduced into Cya of M. intracellulare (Vercellino et al., 2017). Both observations linked the CTE directly to AC activity and emphasized the importance of this element for AC regulation.

Further valuable insight into the structural dynamics of the CTE came from two cases of soluble bacterial AC that had been solved together with their native regulatory domains. Lindner et al. (2017) worked with a subclass IIIb AC from Beggiatoa, which was regulated by an N-terminal blue light sensor domain and could easily be stimulated under laboratory conditions. In the crystal structure of this AC, the CTEs showed the canonical bent conformation. Hydrogen-deuterium exchange experiments with this protein however found the CTE and its upstream coiled coil as the sites of highest deuteration after exposure to light. Remarkably, however, in the presence of a competitive inhibitor, only the N-terminal helix of the CTE showed a deuteration rate similar to that of the upstream coiled coil, whereas deuteration in the C-terminal remained essentially unchanged (Lindner et al., 2017). This indicated that the CTE in the light state had a similar solvent-exposure as the upstream coiled coil. We therefore consider these results as strong support for the proposed structural model of signal transduction in the CTE.

A second interesting insight came from the comparison of bPAC with a recently published structure of the subclass IIIa AC Cya from M. intracellulare. In both structures, the coiled-coil segment upstream of the CTE was sufficiently long to allow for the annotation of the coiled-coil register. Surprisingly, the registers did not match. In Cya, the first core residue in the CTE was in position d, whereas in bPAC it was a (Figure 10B and C). While it is not yet clear to what extent the coiled-coil registers of Cya and bPAC stand representative for their respective subclasses, such fine structural differences provide an attractive hypothesis to explain the sequence divergence between the CTEs of subclasses IIIa and IIIb (Figure 7). Investigating

the functional implications for the differences in the registers, for example as different ground states in a signal transduction mechanism based on axial rotation of the helices, might be an interesting aim for future studies.

3.3 Comprehensive analysis of AC domain architectures

The results summarized in this section were published in Bassler et al. (2018). The full manuscript and a detailed list of my contributions can be found in appendix C.

As I have outlined in the introduction, bacterial ACs occur with a great number of different domain architectures. A great share of this structural diversity originates from the modular recombination of a comparatively small number of input and transducer domains, of which many are also found in other families of signal transduction proteins, including histidine kinases and diguanylate cyclases. Several studies have sought to classify AC domain architectures, but have arguably failed to identify a widespread, functionally meaningful criterion on which to base the classification. A particularly interesting question in this context is how the AC catalytic domains manage to decode regulatory signals from so many, structurally diverse input and transducer domains. Vice versa, this also asks how these input and transducer domains are able to regulate the activity of multiple unrelated output domains. The experimental and bioinformatic characterization of the CTE suggested that this element could be the mediator of diverse upstream signals into a cyclase-specific conformational rearrangement that determines the catalytic activity (Ziegler et al., 2017; see section 3.2), and hence part of the answer to the above questions. Thus, in order to fully understand the role of the CTE, we had to conduct a large-scale analysis of AC domain architectures.

6TM domains define some of the most-abundant domain architectures

I analyzed the domain architectures of 6251 bacterial and 9690 eukaryotic class III nucleotide cyclases in the Uniprot reference proteomes database (release 11/2016; the UniProt Consortium, 2012) using the domain definitions of the SMART (Letunic et al., 2015) and Pfam (Finn et al., 2016) databases and manual annotation. In addition, I identified a small number of ACs from archaea, which appeared to have been acquired by horizontal gene transfer from bacteria. Since their overall number was very low and their sequences and domain architectures were reasonably similar to bacterial ACs, I will not consider archaeal ACs in the following.

As expected, the analysis found a great number of domain architectures. In bacteria, most domain architectures were conserved only between closely related species, while only few architectures were spread across distant lineages. Surprisingly, however, domain architectures with 6TM domains were among the most abundant. In fact, subclass IIIa ACs with a 6TM domain (see section 3.1; Beltz et al., 2016) comprised more than one quarter of all proteins in this subclass and were found in many phyla. Similarly, ACs with a 6TM domain and a HAMP transducer constituted approximately 10% of all domain architectures in subclass IIIb (appendix Figure C.4). Again, this greatly emphasized the importance of 6TM.

The CTE translates AC-unspecific signals into AC-specific rearrangements

Sorting the domain architectures by the presence or absence of a CTE confirmed the previous observation that CTEs were universally present in ACs of subclass IIIa and widespread in ACs of subclass IIIb. Notably, this includes most 6TM ACs except those of the peculiar type associated with a ferredoxin domain. CTEs were absent in a comparatively low number of ACs of subclass IIIb and also from all ACs of subclass IIIc/d. Consequently,subclass IIIb was highly interesting for investigating the role of the CTE. Subclass IIIb ACs with a CTE generally possessed domain architectures that were fundamentally similar or identical to domain architectures found in subclass IIIa. Many of these architectures were characterized by a so-called

small molecule-binding domain, such as PAS, GAF, and CACHE, by a phosphorylation receiver domain, or by the coiled-coil transducers HAMP and the signaling helix (appendix Figures C.3A and C.4A). These domains are frequently found across different signal transduction proteins and are generally positioned along a dimeric coiled-coil backbone. This can, for example, be seen in the structures of the light-activated AC bPAC (Lindner et al., 2017), the chemoreceptor Tsr (Tajima et al., 2011; Ferris et al., 2012), and the histidine kinases VicK and NarQ (Wang et al., 2013; Gushchin et al., 2017). ACs without a CTE, by comparison, were never found with any of these domains and were not predicted to contain coiled coils (appendix Figures C.3B and C.4B). This suggested that the CTE marks the endpoint of a dimeric coiled-coil backbone along which AC-unspecific conformational signals are propagated. The CTE then converts these unspecific signals into an AC-specific structural rearrangement that directly determines AC activity, most likely by means of control over the dimerization of the catalytic domains. It further suggested that the dimeric coiled-coil backbone represents a common interface between many signal transduction proteins that enabled the complex recombination of input and transducer domains. Accordingly, we see a clear analogy between the CTE of ACs and the DHp domain of histidine kinases. Although no comparable domain has yet been described in other coiled-coil signal transduction proteins, such as diguanylate cyclases, phosphatases, and phosphodiesterases, we highly anticipate that more such output domain-specific transducers will be identified in the future.

4 | Conclusion and Outlook

In this dissertation, I have presented a comprehensive bioinformatic classification and characterization of the six-helical transmembrane (6TM) domains of adenylate cyclases (ACs). The classification identified six groups of 6TM domains, each with a distinctive sequence profile. Two closely related groups corresponded to the transmembrane domains of the N-terminal and C-terminal subunits of the eukaryotic membrane-bound ACs (mACs), which are conserved between all animals and are involved in many sensory and developmental processes. Two more groups were found across distant bacterial lineages and represented two of the most abundant architectural types. The outstanding conservation and widespread abundance of these domain types is in stark contrast to our complete ignorance regarding their function.

The experimental work by the Schultz group provided a first indication that some or all AC 6TM domains could be regulatory receptors, possibly for hydrophobic ligands bound inside the membrane. This notion is in agreement with the results of our bioinformatic classification, which suggested a co-adaptation between the 6TM domains and other parts of the proteins, as would be required for the intramolecular propagation of regulatory signals. In parallel, the experimental results pointed out the importance of a short segment on the N-terminus of many AC catalytic domains. Based on a bioinformatic characterization of this segment, we identified the novel cyclase transducer element (CTE). Increasing evidence suggests that this element plays a central role in the decoding of regulatory signals from upstream domains in many class III nucleotide cyclases. Notably, all but one of the identified 6TM domain types were connected to the catalytic domain by a CTE, indicating the propagation of regulatory signals in these proteins. Hence, the identification of this element also reflected back on our original question and provided further support

for the suspected input function of the 6TM domains.

At the same time, our results raised new issues and open questions. The characterization of the CTE suggested that it represented a novel type of output domain-specific transducer element, which converts AC-unspecific signals from the coiled-coil backbone into an AC-specific conformational change. By comparison, other signal transduction proteins with a coiled-coil backbone and similar input and transducer domains should require specific transducers as well. It will be highly interesting to see if more such elements can be identified in the future.

An as-of-yet unanswered question concerns the relationship between the various types of 6TM domains. In my 2016 analysis, the low-level similarities between five of the six groups (i.e. excluding the ferredoxin-associated group) tentatively suggested a homologous relationship between these domains, but did not allow for a definite statement. In the meantime, however, the discovery of the CTE and its role in the recombination of domains between signal transduction proteins suggested that the closest homologs of AC 6TM domains may well be found outside the AC family. In that case, the AC 6TM domains could be part of a much larger superfamily of 6TM input domains, similar to the small molecule-binding domain superfamily of soluble receptor domains.

However, the most important aim for future work is clearly the identification of the natural ligands of the 6TM domains. We have identified an increasing amount of circumstantial evidence, which suggested that not only the bacterial, but also the eukaryotic 6TM may be cryptic regulatory receptors; the final prove of this hypothesis, however, can only be provided by the identification of the natural ligands. Especially the discovery of ligands of the mammalian mACs will have tremendous pharmacological impact and could significantly expand our understanding of cellular signaling. I am extremely proud that our results have spurred several new initiatives to investigate the structures, functions, and putative ligands of 6TM domains. I look forward to see exciting results in this field in the near future.

References

Acin-Perez, Rebeca et al. (Mar. 2009). "Cyclic AMP produced inside mitochondria regulates oxidative phosphorylation." In: Cell metabolism 9.3, pp. 265–76. ISSN: 1932-7420. DOI: 10.1016/j.cmet.2009.01.012.

Anantharaman, Vivek and L. Aravind (Nov. 2000). "Cache - a signaling domain common to animal Ca(2+)-channel subunits and a class of prokaryotic chemotaxis receptors." In: Trends in biochemical sciences 25.11, pp. 535–7. ISSN: 0968-0004. DOI: 10.1016/S0968-0004(00)01672-8.

Anantharaman, Vivek, S Balaji, and L Aravind (Sept. 2006). "The signaling helix: a common functional theme in diverse signaling proteins." In: Biology direct 1, p. 25. ISSN: 1745-6150. DOI: 10.1186/1745-6150-1-25.

Anantharaman, Vivek, Eugene V Koonin, and L Aravind (Apr. 2001). "Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains." In: Journal of molecular biology 307.5, pp. 1271–92. ISSN: 0022-2836. DOI: 10.1006/jmbi.2001.4508.

Appleman, J. Alex, Li-Ling Chen, and Valley Stewart (Aug. 2003). "Probing conservation of HAMP linker structure and signal transduction mechanism through analysis of hybrid sensor kinases." In: Journal of bacteriology 185.16, pp. 4872–82. ISSN: 0021-9193. DOI: 10.1128/JB.185.16.4872-4882.2003.

Artymiuk, P J et al. (July 1997). "A polymerase I palm in adenylyl cyclase?" In: Nature 388.6637, pp. 33–4. ISSN: 0028-0836. DOI: 10.1038/40310.

Baker, David A (Sept. 2004). "Adenylyl and guanylyl cyclases from the malaria parasite Plasmodium falciparum." In: IUBMB life 56.9, pp. 535–40. ISSN: 1521-6543. DOI: 10.1080/15216540400013937.

Bassler, Jens, Joachim E. Schultz, and Andrei N. Lupas (June 2018). "Adenylate cyclases: Receivers, transducers, and generators of signals". In: Cellular Signalling 46, pp. 135–144. ISSN: 08986568. DOI: 10.1016/j.cellsig.2018.03.002.

Beltz, Stephanie, Jens Bassler, and Joachim E Schultz (2016). "Regulation by the quorum sensor from Vibrio indicates a receptor function for the membrane anchors of adenylate cyclases". In: eLife 5, pp. 1–23. ISSN: 2050-084X. DOI: 10.7554/eLife.13098.

Bos, Johannes L. (Dec. 2006). "Epac proteins: multi-purpose cAMP targets." In: Trends in biochemical sciences 31.12, pp. 680–6. ISSN: 0968-0004. DOI: 10.1016/j.tibs.2006.10.002.

Cann, Martin (Apr. 2007). "A subset of GAF domains are evolutionarily conserved sodium sensors." In: Molecular microbiology 64.2, pp. 461–72. ISSN: 0950-382X. DOI: 10.1111/j.1365-2958.2007.05669.x.

Chen, J et al. (May 1995). "A region of adenylyl cyclase 2 critical for regulation by G protein beta gamma subunits." In: Science 268.5214, pp. 1166–9. ISSN: 0036-8075.

Colledge, Marcie and John D Scott (June 1999). "AKAPs: from structure to function." In: Trends in cell biology 9.6, pp. 216–21. ISSN: 0962-8924. DOI: 10.1016/S0962-8924(99)01558-5.

Cooper, Dermot M F and Andrew J. Crossthwaite (Aug. 2006). "Higher-order organization and regulation of adenylyl cyclases." In: Trends in pharmacological sciences 27.8, pp. 426–31. ISSN: 0165-6147. DOI: 10.1016/j.tips.2006.06.002.

Cotta, M. A., T. R. Whitehead, and M. B. Wheeler (July 1998). "Identification of a novel adenylate cyclase in the ruminal anaerobe, Prevotella ruminicola D31d." In: FEMS microbiology letters 164.2, pp. 257–60. ISSN: 0378-1097. DOI: 10.1016/S0378-1097(98)00221-3.

Danchin, Antoine (1993). "Phylogeny of adenylyl cyclases." In: Advances in second messenger and phosphoprotein research 27, pp. 109–62. ISSN: 1040-7952.

Dessauer, Carmen W et al. (Apr. 2017). "International Union of Basic and Clinical Pharmacology. CI. Structures and Small Molecule Modulators of Mammalian Adenylyl Cyclases". In: Pharmacological Reviews 69.2. Ed. by Eliot H. Ohlstein, pp. 93–139. ISSN: 0031-6997. DOI: 10.1124/pr.116.013078.

Devreotes, Peter (Sept. 1989). "Dictyostelium discoideum: a model system for cell-cell interactions in development". In: Science 245.4922, pp. 1054–1058. ISSN: 0036-8075. DOI: 10.1126/science.2672337.

Dunin-Horkawicz, Stanislaw and Andrei N. Lupas (Apr. 2010). "Comprehensive analysis of HAMP domains: implications for transmembrane signal transduction." In: Journal

of molecular biology 397.5, pp. 1156–74. ISSN: 1089-8638. DOI: 10.1016/j.jmb.2010.02.031.

Etzl, Stefan et al. (Apr. 2018). "Structure-guided design and functional characterization of an artificial red light-regulated guanylate/adenylate cyclase for optogenetic applications." In: The Journal of biological chemistry. ISSN: 1083-351X. DOI: 10.1074/jbc.RA118.003069.

Ferris, Hedda U. et al. (Mar. 2011). "The mechanisms of HAMP-mediated signaling in transmembrane receptors." In: Structure 19.3, pp. 378–85. ISSN: 1878-4186. DOI: 10.1016/j.str.2011.01.006.

Ferris, Hedda U. et al. (Jan. 2012). "Mechanism of regulation of receptor histidine kinases." In: Structure 20.1, pp. 56–66. ISSN: 1878-4186. DOI: 10.1016/j.str.2011.11.014.

Finn, Robert D et al. (Jan. 2016). "The Pfam protein families database: towards a more sustainable future." In: Nucleic acids research 44.D1, pp. D279–85. ISSN: 1362-4962. DOI: 10.1093/nar/gkv1344.

Frickey, Tancred and Andrei Lupas (Dec. 2004). "CLANS: a Java application for visualizing protein families based on pairwise similarity." In: Bioinformatics 20.18, pp. 3702–4. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/bth444.

Fushimi, Keiji et al. (May 2017). "Distinctive Properties of Dark Reversion Kinetics between Two Red/Green-Type Cyanobacteriochromes and their Application in the Photoregulation of cAMP Synthesis." In: Photochemistry and photobiology 93.3, pp. 681–691. ISSN: 1751-1097. DOI: 10.1111/php.12732.

Gancedo, Juana M. (Aug. 2013). "Biological roles of cAMP: variations on a theme in the different kingdoms of life." In: Biological reviews of the Cambridge Philosophical Society 88.3, pp. 645–68. ISSN: 1469-185X. DOI: 10.1111/brv.12020.

Gehring, Chris and Ilona S. Turek (2017). "Cyclic Nucleotide Monophosphates and Their Cyclases in Plant Signaling." In: Frontiers in plant science 8.October, p. 1704. ISSN: 1664-462X. DOI: 10.3389/fpls.2017.01704.

Gersch, D, W Römer, and H Krügel (June 1979). "Inverse regulation of spore germination and growth by cyclic AMP in Streptomyces hygroscopicus." In: Experientia 35.6, p. 749. ISSN: 0014-4754.

Gilman, Alfred G (1987). "G proteins: transducers of receptor-generated signals." In: Annual review of biochemistry 56, pp. 615–49. ISSN: 0066-4154. DOI: 10.1146/annurev.bi.56.070187.003151.

Guo, Ying Lan et al. (July 2001). "Adenylyl cyclase Rv1625c of Mycobacterium tuberculosis: a progenitor of mammalian adenylyl cyclases." In: The EMBO journal 20.14, pp. 3667–75. ISSN: 0261-4189. DOI: 10.1093/emboj/20.14.3667.

Gushchin, Ivan et al. (2017). "Mechanism of transmembrane signaling by sensor histidine kinases." In: Science 356.6342. ISSN: 1095-9203. DOI: 10.1126/science.aah6345.

Hanks, Steven K and Tony Hunter (May 1995). "The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification." In: FASEB J 9.8, pp. 576–96. ISSN: 0892-6638. DOI: 10.1096/fasebj.9.8.7768349.

Hefti, Marco H. et al. (Mar. 2004). "The PAS fold. A redefinition of the PAS domain based upon structural prediction." In: European journal of biochemistry 271.6, pp. 1198–208. ISSN: 0014-2956. DOI: 10.1111/j.1432-1033.2004.04023.x.

Henry, Jonathan T and Sean Crosson (2011). "Ligand-binding PAS domains in a genomic, cellular, and structural context." In: Annual review of microbiology 65, pp. 261–86. ISSN: 1545-3251. DOI: 10.1146/annurev-micro-121809-151631.

Hewlett, E. and J. Wolff (Aug. 1976). "Soluble adenylate cyclase from the culture medium of Bordetella pertussis: purification and characterization." In: Journal of bacteriology 127.2, pp. 890–8. ISSN: 0021-9193.

Higgins, Douglas A. et al. (Dec. 2007). "The major Vibrio cholerae autoinducer and its role in virulence factor production." In: Nature 450.7171, pp. 883–6. ISSN: 1476-4687. DOI: 10.1038/nature06284.

Hulko, Michael et al. (Sept. 2006). "The HAMP domain structure implies helix rotation in transmembrane signaling." In: Cell 126.5, pp. 929–40. ISSN: 0092-8674. DOI: 10.1016/j.cell.2006.06.058.

Jones, D T (1999). "Protein secondary structure prediction based on position-specific scoring matrices." In: Journal of molecular biology 292, pp. 195–202. ISSN: 0022-2836. DOI: 10.1006/jmbi.1999.3091.

Kanacher, Tobias et al. (July 2002). "A GAF-domain-regulated adenylyl cyclase from Anabaena is a self-activating cAMP switch." In: The EMBO journal 21.14, pp. 3672–80. ISSN: 0261-4189. DOI: 10.1093/emboj/cdf375.

Kanchan, Kajal et al. (Jan. 2010). "Transmembrane signaling in chimeras of the Escherichia coli aspartate and serine chemotaxis receptors and bacterial class III adenylyl cyclases." In: The Journal of biological chemistry 285.3, pp. 2090–9. ISSN: 1083-351X. DOI: 10.1074/jbc.M109.051698.

Karniol, Baruch et al. (Nov. 2005). “Phylogenetic analysis of the phytochrome superfamily reveals distinct microbial subfamilies of photoreceptors.” In: The Biochemical journal 392.Pt 1, pp. 103–16. ISSN: 1470-8728. DOI: 10.1042/BJ20050826.

Kasahara, Masahiro and Masayuki Ohmori (May 1999). “Activation of a cyanobacterial adenylate cyclase, CyaC, by autophosphorylation and a subsequent phosphotransfer reaction.” In: The Journal of biological chemistry 274.21, pp. 15167–72. ISSN: 0021-9258. DOI: 10.1074/jbc.274.21.15167.

Kaupp, U Benjamin and Reinhard Seifert (July 2002). “Cyclic nucleotide-gated ion channels.” In: Physiological reviews 82.3, pp. 769–824. ISSN: 0031-9333. DOI: 10.1152/physrev.00008.2002.

Kleinboelting, Silke et al. (Mar. 2014). “Crystal structures of human soluble adenylyl cyclase reveal mechanisms of catalysis and of its activation through bicarbonate.” In: Proceedings of the National Academy of Sciences of the United States of America 111.10, pp. 3727–32. ISSN: 1091-6490. DOI: 10.1073/pnas.1322778111.

Krupinski, John et al. (June 1989). “Adenylyl cyclase amino acid sequence: possible channel- or transporter-like structure”. In: Science 244.4912, pp. 1558–1564. ISSN: 0036-8075. DOI: 10.1126/science.2472670.

Leppla, S H (May 1982). “Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells.” In: Proceedings of the National Academy of Sciences of the United States of America 79.10, pp. 3162–6. ISSN: 0027-8424. DOI: 10.1073/pnas.79.10.3162.

Letunic, I., T. Doerks, and P. Bork (2015). “SMART: recent updates, new developments and status in 2015”. In: Nucleic Acids Research 43.D1, pp. D257–D260. ISSN: 0305-1048. DOI: 10.1093/nar/gku949.

Linder, Jürgen U. (Jan. 2010). “cGMP production in bacteria.” In: Molecular and cellular biochemistry 334.1-2, pp. 215–9. ISSN: 1573-4919. DOI: 10.1007/s11010-009-0321-0.

Linder, Jürgen U. and Joachim E. Schultz (Dec. 2003). “The class III adenylyl cyclases: multi-purpose signalling modules.” In: Cellular signalling 15.12, pp. 1081–9. ISSN: 0898-6568. DOI: 10.1016/S0898-6568(03)00130-X.

Linder, Jürgen U. et al. (Aug. 1999). “Guanylyl cyclases with the topology of mammalian adenylyl cyclases and an N-terminal P-type ATPase-like domain in Paramecium, Tetrahy-

mena and Plasmodium." In: The EMBO journal 18.15, pp. 4222–32. ISSN: 0261-4189. DOI: 10.1093/emboj/18.15.4222.

Linder, Jürgen U, Anita Schultz, and Joachim E Schultz (May 2002). "Adenylyl cyclase Rv1264 from Mycobacterium tuberculosis has an autoinhibitory N-terminal domain." In: The Journal of biological chemistry 277.18, pp. 15271–6. ISSN: 0021-9258. DOI: 10.1074/jbc.M200235200.

Lindner, Robert et al. (2017). "Photoactivation Mechanism of a Bacterial Light-Regulated Adenylyl Cyclase". In: Journal of Molecular Biology 429.9, pp. 1336–1351. ISSN: 10898638. DOI: 10.1016/j.jmb.2017.03.020.

Ma, Xiaolei, Annie Beuve, and Focco van den Akker (Jan. 2010). "Crystal structure of the signaling helix coiled-coil domain of the beta1 subunit of the soluble guanylyl cyclase." In: BMC structural biology 10.1, p. 2. ISSN: 1472-6807. DOI: 10.1186/1472-6807-10-2.

Marden, Jeremiah N. et al. (Feb. 2011). "Cyclic GMP controls Rhodospirillum centenum cyst development." In: Molecular microbiology 79.3, pp. 600–15. ISSN: 1365-2958. DOI: 10.1111/j.1365-2958.2010.07513.x.

McDonough, Kathleen A. and Ana Rodriguez (Nov. 2011). "The myriad roles of cyclic AMP in microbial pathogens: from signal to sword." In: Nature reviews. Microbiology 10.1, pp. 27–38. ISSN: 1740-1534. DOI: 10.1038/nrmicro2688.

Mondéjar, Laura García et al. (Jan. 2012). "HAMP domain-mediated signal transduction probed with a mycobacterial adenylyl cyclase as a reporter." In: The Journal of biological chemistry 287.2, pp. 1022–31. ISSN: 1083-351X. DOI: 10.1074/jbc.M111.284067.

Mougel, Christophe and Igor B. Zhulin (Oct. 2001). "CHASE: an extracellular sensing domain common to transmembrane receptors from prokaryotes, lower eukaryotes and plants." In: Trends in biochemical sciences 26.10, pp. 582–4. ISSN: 0968-0004. DOI: 10.1016/S0968-0004(01)01969-7.

Ng, Wai-Leung and Bonnie L. Bassler (Dec. 2009). "Bacterial Quorum-Sensing Network Architectures". In: Annual Review of Genetics 43.1, pp. 197–222. ISSN: 0066-4197. DOI: 10.1146/annurev-genet-102108-134304.

Nikolskaya, Anastsia N et al. (2003). "MASE1 and MASE2: two novel integral membrane sensory domains." In: Journal of molecular microbiology and biotechnology 5.1, pp. 11–6. ISSN: 1464-1801. DOI: 10.1159/000068720.

Nixon, B. T., C. W. Ronson, and F. M. Ausubel (Oct. 1986). "Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC." In: Proceedings of the National Academy of Sciences of the United States of America 83.20, pp. 7850–4. ISSN: 0027-8424. DOI: 10.1073/pnas.83.20.7850.

Ohki, Mio et al. (2016). "Structural insight into photoactivation of an adenylate cyclase from a photosynthetic cyanobacterium." In: Proceedings of the National Academy of Sciences of the United States of America 113.24, pp. 6659–64. ISSN: 1091-6490. DOI: 10.1073/pnas.1517520113.

Ponting, C. P. et al. (June 1999). "Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer." In: Journal of molecular biology 289.4, pp. 729–45. ISSN: 0022-2836. DOI: 10.1006/jmbi.1999.2827.

Reddy, R et al. (June 1995). "Voltage-sensitive adenylyl cyclase activity in cultured neurons. A calcium-independent phenomenon." In: The Journal of biological chemistry 270.24, pp. 14340–6. ISSN: 0021-9258.

Remmert, Michael et al. (Feb. 2012). "HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment." In: Nature methods 9.2, pp. 173–5. ISSN: 1548-7105. DOI: 10.1038/nmeth.1818.

Roelofs, Jeroen et al. (Mar. 2001). "Guanylate cyclase in Dictyostelium discoideum with the topology of mammalian adenylate cyclase." In: The Biochemical journal 354.Pt 3, pp. 697–706. ISSN: 0264-6021.

Ryu, Min-Hyung et al. (May 2015). "Identification of bacterial guanylate cyclases". In: Proteins: Structure, Function, and Bioinformatics 83.5, pp. 799–804. ISSN: 08873585. DOI: 10.1002/prot.24769.

Sadana, Rachna and Carmen W. Dessauer (2009). "Physiological roles for G protein-regulated adenylyl cyclase isoforms: Insights from knockout and overexpression studies". In: NeuroSignals 17.1, pp. 5–22. ISSN: 1424862X. DOI: 10.1159/000166277.

Schultz, Joachim E. and Janani Natarajan (Nov. 2013). "Regulated unfolding: a basic principle of intraprotein signaling in modular proteins." In: Trends in biochemical sciences 38.11, pp. 538–45. ISSN: 0968-0004. DOI: 10.1016/j.tibs.2013.08.005.

Shenoy, Avinash R and Sandhya S. Visweswariah (Mar. 2004). "Class III nucleotide cyclases in bacteria and archaebacteria: lineage-specific expansion of adenylyl cyclases

and a dearth of guanylyl cyclases." In: FEBS letters 561.1-3, pp. 11–21. ISSN: 0014-5793. DOI: 10.1016/S0014-5793(04)00128-0.

Sinha, Sangita C et al. (Feb. 2005). "Origin of asymmetry in adenylyl cyclases: structures of Mycobacterium tuberculosis Rv1900c." In: The EMBO journal 24.4, pp. 663–73. ISSN: 0261-4189. DOI: 10.1038/sj.emboj.7600573.

Sismeiro, Odile, Pascale Trotot, and Antoine Danchin (1998). "Aeromonas hydrophila Adenylyl Cyclase 2: a New Class of Adenylyl Cyclases". In: Journal of Bacteriology 180.13, p. 3339.

Spirig, Thomas et al. (June 2008). "The Legionella autoinducer synthase LqsA produces an alpha-hydroxyketone signaling molecule." In: The Journal of biological chemistry 283.26, pp. 18113–23. ISSN: 0021-9258. DOI: 10.1074/jbc.M801929200.

Stierl, Manuela et al. (Jan. 2011). "Light modulation of cellular cAMP by a small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium Beggiatoa." In: The Journal of biological chemistry 286.2, pp. 1181–8. ISSN: 1083-351X. DOI: 10.1074/jbc.M110.185496.

Stock, Ann M, Victoria L Robinson, and Paul N Goudreau (2000). "Two-component signal transduction." In: Annual review of biochemistry 69, pp. 183–215. ISSN: 0066-4154. DOI: 10.1146/annurev.biochem.69.1.183.

Süsstrunk, Urs et al. (Oct. 1998). "Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in Streptomyces coelicolor." In: Molecular microbiology 30.1, pp. 33–46. ISSN: 0950-382X. DOI: 10.1046/j.1365-2958.1998.01033.x.

Tajima, Hirotaka et al. (Dec. 2011). "Ligand specificity determined by differentially arranged common ligand-binding residues in bacterial amino acid chemoreceptors Tsr and Tar." In: The Journal of biological chemistry 286.49, pp. 42200–10. ISSN: 1083-351X. DOI: 10.1074/jbc.M111.221887.

Téllez-Sosa, Juan et al. (July 2002). "The Rhizobium etli cyaC product: characterization of a novel adenylate cyclase class." In: Journal of bacteriology 184.13, pp. 3560–8. ISSN: 0021-9193. DOI: 10.1128/JB.184.13.3560.

Tesmer, J J and S R Sprang (Dec. 1998). "The structure, catalytic mechanism and regulation of adenylyl cyclase." In: Current opinion in structural biology 8.6, pp. 713–9. ISSN: 0959-440X.

Tesmer, J J et al. (1997). “Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS.” In: Science 278.December, pp. 1907–1916. ISSN: 0036-8075. DOI: 10.1126/science.278.5345.1907.

Tesmer, J. J. et al. (July 1999). “Two-metal-Ion catalysis in adenylyl cyclase.” In: Science 285.5428, pp. 756–60. ISSN: 0036-8075. DOI: 10.1126/science.285.5428.756.

Tews, Ivo et al. (May 2005). “The structure of a pH-sensing mycobacterial adenylyl cyclase holoenzyme.” In: Science 308.5724, pp. 1020–3. ISSN: 1095-9203. DOI: 10.1126/science.1107642.

Tiaden, André and Hubert Hilbi (2012). “alpha-hydroxyketone synthesis and sensing by Legionella and Vibrio”. In: Sensors 12.3, pp. 2899–2919. ISSN: 14248220. DOI: 10.3390/s120302899.

Turnham, Rigney E. and John D. Scott (Feb. 2016). “Protein kinase A catalytic subunit isoform PRKACA; History, function and physiology.” In: Gene 577.2, pp. 101–8. ISSN: 1879-0038. DOI: 10.1016/j.gene.2015.11.052.

Ulrich, Luke E, Eugene V Koonin, and Igor B Zhulin (Feb. 2005). “One-component systems dominate signal transduction in prokaryotes”. In: Trends in Microbiology 13.2, pp. 52–56. ISSN: 0966842X. DOI: 10.1016/j.tim.2004.12.006.

UniProt Consortium (Jan. 2012). “Reorganizing the protein space at the Universal Protein Resource (UniProt).” In: Nucleic acids research 40.Database issue, pp. D71–5. ISSN: 1362-4962. DOI: 10.1093/nar/gkr981.

Upadhyay, Amit A et al. (Apr. 2016). “Cache Domains That are Homologous to, but Different from PAS Domains Comprise the Largest Superfamily of Extracellular Sensors in Prokaryotes.” In: PLoS computational biology 12.4, e1004862. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1004862.

Utsumi, R et al. (Sept. 1989). “Activation of bacterial porin gene expression by a chimeric signal transducer in response to aspartate.” In: Science 245.4923, pp. 1246–9. ISSN: 0036-8075.

Vercellino, Irene et al. (Nov. 2017). “Role of the nucleotidyl cyclase helical domain in catalytically active dimer formation.” In: Proceedings of the National Academy of Sciences of the United States of America 114.46, E9821–E9828. ISSN: 1091-6490. DOI: 10.1073/pnas.1712621114.

Vorherr, T et al. (June 1993). “The calmodulin binding domain of nitric oxide synthase and adenylyl cyclase.” In: Biochemistry 32.23, pp. 6081–8. ISSN: 0006-2960.

Wang, Chen et al. (2013). "Mechanistic insights revealed by the crystal structure of a histidine kinase with signal transducer and sensor domains." In: PLoS biology 11.2, e1001493. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.1001493.

Weng, Gezhi et al. (Oct. 1996). "Gbeta subunit interacts with a peptide encoding region 956-982 of adenylyl cyclase 2. Cross-linking of the peptide to free Gbetagamma but not the heterotrimer." In: The Journal of biological chemistry 271.43, pp. 26445–8. ISSN: 0021-9258. DOI: 10.1074/jbc.271.43.26445.

Winkler, Karin, Anita Schultz, and Joachim E. Schultz (May 2012). "The S-helix determines the signal in a Tsr receptor/adenylyl cyclase reporter." In: The Journal of biological chemistry 287.19, pp. 15479–88. ISSN: 1083-351X. DOI: 10.1074/jbc.M112.348409.

Wu, Zhiliang, Scott T. Wong, and D R Storms (Nov. 1993). "Modification of the calcium and calmodulin sensitivity of the type I adenylyl cyclase by mutagenesis of its calmodulin binding domain." In: The Journal of biological chemistry 268.32, pp. 23766–8. ISSN: 0021-9258.

Yan, Shui Zhong et al. (May 1997). "The conserved asparagine and arginine are essential for catalysis of mammalian adenylyl cyclase." In: The Journal of biological chemistry 272.19, pp. 12342–9. ISSN: 0021-9258. DOI: 10.1074/jbc.272.19.12342.

Zhang, Gongyi et al. (Mar. 1997). "Structure of the adenylyl cyclase catalytic core." In: Nature 386.6622, pp. 247–53. ISSN: 0028-0836. DOI: 10.1038/386247a0.

Zhulin, Igor B., Anastasia N. Nikolskaya, and Michael Y. Galperin (Jan. 2003). "Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea." In: Journal of bacteriology 185.1, pp. 285–94. ISSN: 0021-9193. DOI: 10.1128/JB.185.1.285-294.2003.

Ziegler, Miriam et al. (2017). "Characterization of a novel signal transducer element intrinsic to class IIIa/b adenylate cyclases and guanylate cyclases." In: The FEBS journal 284.8, pp. 1204–1217. ISSN: 1742-4658. DOI: 10.1111/febs.14047.

Appendix

A Manuscript Beltz et al. (2016) eLIFE

Beltz S., *Bassler J.*, and Schultz J. E.

Regulation by the quorum sensor from Vibrio indicates a receptor function for the membrane anchors of adenylate cyclases

in: eLIFE 5 (2016)

Personal contribution:

Conception of the bioinformatic analysis

Acquisition of data shown in figures 7 and 7.S1

Analysis and interpretation of this data (with JES)

Drafting and revising the manuscript (with all authors)

eLIFE elifesciences.org

RESEARCH ARTICLE

Regulation by the quorum sensor from Vibrio indicates a receptor function for the membrane anchors of adenylate cyclases

Stephanie Beltz[1], Jens Bassler[2], Joachim E Schultz[1]*

[1]Pharmazeutisches Institut der Universität Tübingen, Tübingen, Germany; [2]Max-Planck-Institut für Entwicklungsbiologie, Tübingen, Germany

Abstract Adenylate cyclases convert intra- and extracellular stimuli into a second messenger cAMP signal. Many bacterial and most eukaryotic ACs possess membrane anchors with six transmembrane spans. We replaced the anchor of the AC Rv1625c by the quorum-sensing receptor from *Vibrio harveyi* which has an identical 6TM design and obtained an active, membrane-anchored AC. We show that a canonical class III AC is ligand-regulated *in vitro* and *in vivo*. At 10 μM, the cholera-autoinducer CAI-1 stimulates activity 4.8-fold. A sequence based clustering of membrane domains of class III ACs and quorum-sensing receptors established six groups of potential structural and functional similarities. The data support the notion that 6TM AC membrane domains may operate as receptors which directly regulate AC activity as opposed and in addition to the indirect regulation by GPCRs in eukaryotic congeners. This adds a completely novel dimension of potential AC regulation in bacteria and vertebrates.
DOI: 10.7554/eLife.13098.001

***For correspondence:** joachim.schultz@uni-tuebingen.de

Competing interests: The author declares that no competing interests exist.

Funding: See page 14

Received: 17 November 2015
Accepted: 26 February 2016
Published: 27 February 2016

Reviewing editor: Michael A Marletta, University of California, Berkeley, United States

Introduction

In 1958 Sutherland and Rall reported the structure of a second messenger, cyclic 3',5'-adenosine monophosphate (cAMP) which was generated upon incubation of a liver extract with the first messengers epinephrine or glucagon (***Sutherland and Rall, 1958***). Since then cAMP has been demonstrated to be a universal second messenger translating a variety of extracellular stimuli into a uniform intracellular chemical signal. The enzymes responsible for biosynthesis of cAMP from ATP, adenylate cyclases (ACs), have been biochemically and genetically identified in most bacterial and eukaryotic cells (***Khandelwal and Hamilton, 1971***; ***Linder and Schultz, 2003***; ***2008***). To date, sequencing has identified six classes of ACs. The small-sized class I (enterobacterial ACs), class II (toxin class) and the minor classes IV-VI are restricted to bacteria (***Bârzu and Danchin, 1994***; ***Linder and Schultz, 2003***). The class III ACs are ubiquitous, albeit with differing domain architectures (***Linder and Schultz, 2008***). The catalytic domains share sequence and structural similarities, yet minor, characteristic sequence peculiarities have resulted in a division into four subclasses, a-d (***Linder and Schultz, 2008***; ***Tesmer et al., 1997***; ***Tews et al., 2005***). A fundamental difference between bacterial and eukaryotic ACs is that the former are monomers which require homodimerization for activity. The mammalian congeners, exclusively class IIIa, present themselves as pseudoheterodimers composed of two concatenated 'bacterial' monomers with slightly diverged, yet complementary domains (***Guo et al., 2001***). Accordingly, mammalian class III ACs are anchored to the membrane by two putative 6TM bundles, one in each of the concatenated repeats. Our knowledge about regulation of bacterial class III ACs is limited. Apart from a few soluble ACs which appear to be regulated by carbon dioxide or pH near to nothing is known (***Kleinboelting et al.,***

eLife digest Cells are surrounded by a membrane that separates the inside of the cell from the external environment. To communicate information across the cell membrane, cells often employ a relay system. In this system, receptor proteins on the surface of the cells sense information about the environment and trigger the production of a chemical signal inside the cell.

Certain receptors activate enzymes called adenylate cyclases, which reside just inside the cell, to produce a chemical signal. In some human and bacterial adenylate cyclases, about 40% of the protein is anchored in the membrane, far more than is necessary to hold the protein in place. It is therefore possible that this "membrane anchor" region plays an additional role, perhaps even detecting external signals.

A "quorum sensing" receptor protein that was recently discovered embedded in the membrane of a species of marine bacteria called *Vibrio harveyi* has a similar structure to the membrane anchor of adenylate cyclases. Beltz et al. have now replaced the adenylate cyclase membrane anchor with a *V. harveyi* receptor. This produced a hybrid protein that could both receive and translate signals from the membrane receptor.

A computational analysis of the membrane anchors of adenylate cyclases showed that they have striking similarities to quorum sensors. Furthermore, the membrane anchors of different types of adenylate cyclase have diverse structures that may have helped the cyclases to adapt to different environments and biological requirements.

Overall, Beltz et al.'s results suggest the adenylate cyclase membrane anchor is a new type of cell surface receptor. In the future it will be important to identify the environmental signal that activates adenylate cyclases, both in bacteria and mammals.

DOI: 10.7554/eLife.13098.002

2014; ***Steegborn et al., 2005***; ***Tews et al., 2005***). The established regulation of the nine mammalian, membrane-delimited AC isoforms is indirect. Stimulation of G-protein-coupled receptors (GPCRs) by extracellular ligands releases Gsα intracellularly which binds to ACs and activates. A potentially direct ligand-regulation of class III ACs via their large membrane anchors remains a genuine possibility.

The 6TM membrane anchors of bacterial ACs are obviously structural analogs of the 6TM bundles in mammalian ACs. In the past, we replaced the membrane anchor of the mycobacterial AC Rv3645 by the *E. coli* chemotaxis receptors for serine, Tsr, and aspartate, Tar (***Kanchan et al., 2010***). Tsr/Tar and Rv3645 have a signal-transducing HAMP domain in common and both require dimerization. The chimeras were regulated *in vitro* and *in vivo* by serine or aspartate (***Kanchan et al., 2010***; ***Mondéjar et al., 2012***; ***Winkler et al., 2012***), i.e. a 2TM receptor, Tsr or Tar, with an extensive periplasmic ligand-binding domain replaced a 6TM AC membrane anchor which lacks periplasmic loops. The data demonstrated that in principle direct regulation of a class III AC via an extracellular ligand is a possibility. The question whether also a 6TM receptor might directly regulate a 6TM AC remained open. This question is addressed here.

Membrane anchors with 6TMs are present in many proteins. Often they have short transmembrane-spanning α-helices and short connecting loops, e.g. in bacterial and mammalian ACs (***Krupinski et al., 1989***; ***Linder and Schultz, 2003***), in the cytochrome subunits of succinate dehydrogenases and fumarate reductases (***Hederstedt, 1998***; ***Yankovskaya et al., 2003***), ABC transporters (***Chang and Roth, 2001***), in bacterial HdeD proteins (***Mates et al., 2007***), **s**ix **t**ransmembrane **e**pithelial **a**ntigen of the **p**rostate (STEAP, ***Kleven et al., 2015***), or quorum-sensing (QS) receptors from *Vibrio* and *Legionella* which have His-kinases as cytosolic effectors (***Ng and Bassler, 2009***). For the latter lipophilic ligands have been identified (***Ng et al., 2011***; ***2010***). This has opened the opportunity to replace the 6TM anchor of the mycobacterial class IIIa AC Rv1625c which is considered to be an ancestral form of mammalian ACs (***Guo et al., 2001***), by the prototypically identical 6TM QS-receptor CqsS from *V. harveyi* (***Figure 1***). Here, as a proof of principle, we demonstrate that a 6TM receptor not only substitutes a membrane-anchoring function, but also confers direct regulation of a class IIIa AC via an extracellular ligand with nanomolar potency. Taken together with a bioinformatic analysis the data indicate that the 6TM AC membrane anchors

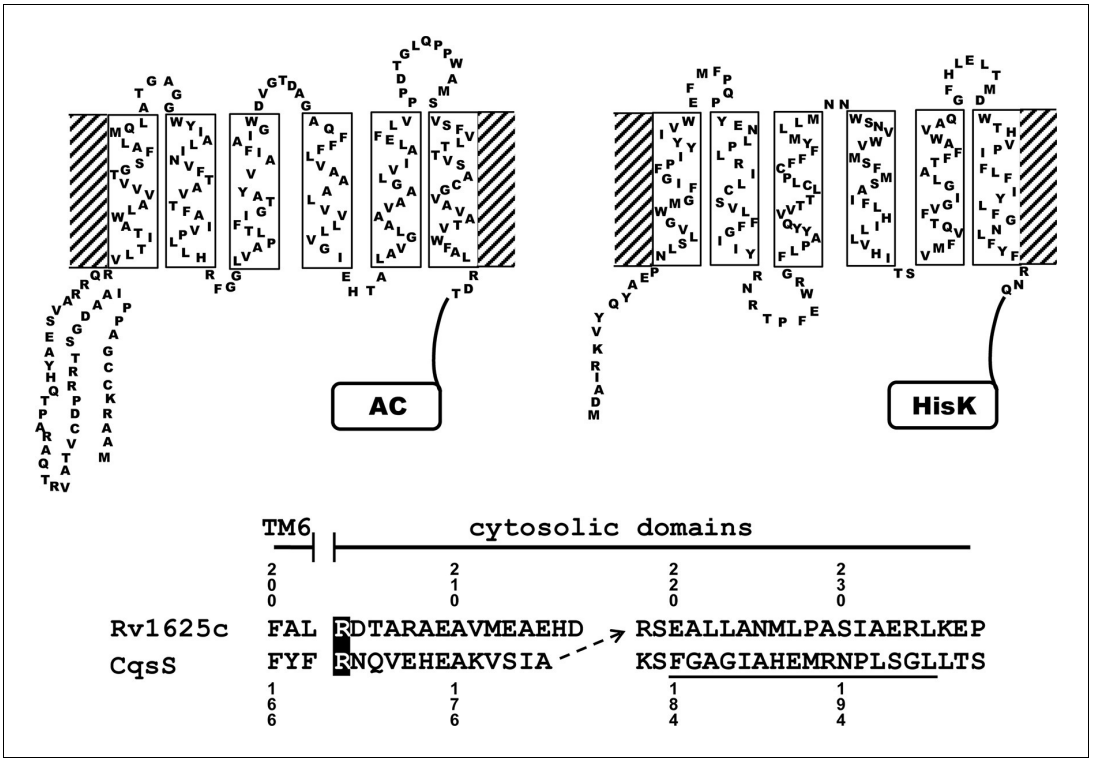


Figure 1. Two-dimensional models of the canonical class IIIa adenylate cyclase Rv1625c from *M. tuberculosis* (left) and the quorum-sensing receptor from *V. harveyi* (right) in the membrane. Both proteins require dimerization to be catalytically active. The alignment below covers the amino acid sequences at the exit of TM6 of both proteins. The most efficient functional linkage of the CqsS receptor to the catalytic domain of Rv1625c is indicated by an arrow. The H-Box of the histidine-kinase domain is underlined. The numbering for CqsS and Rv1625c is indicated above and below the respective sequences.

DOI: 10.7554/eLife.13098.003

probably have an additional receptor function. This is a first and decisive step to add a novel dimension of direct regulation of class IIIa AC activities.

Results

The quorum-sensing receptor CqsS from *V. harveyi* regulates the class IIIa AC Rv1625c

Many bacterial and the nine membrane-bound mammalian class IIIa ACs possess 6TM modules as membrane anchors (***Guo et al., 2001***; ***Linder and Schultz, 2003***; ***2008***). They comprise >40% of the proteins. A function of these membrane domains beyond fixation in the membrane is unknown. Recently we became aware of the QS-receptors from *Vibrio* and *Legionella*, CqsS and LqsS, which feature a membrane anchor of an essentially identical design as the aforementioned ACs, i.e. minimal-length α-helices and short connecting linkers which presumably severely restrict conformational possibilities (see ***Figure 1*** for a 2D representation; ***Ng et al., 2010***; ***Tiaden and Hilbi, 2012***). For both QS-receptors highly lipophilic ligands have been identified such as CAI-1, the ***C**holerae* **A**uto**I**nducer-**1**, (S)-3-hydroxy-tridecan-4-one and LAI-1, the ***L**egionella* **A**uto**I**nducer-**1**, (S)-3-hydroxy-pentadecan-4-one (***Ng et al., 2010***; ***Spirig et al., 2008***). The QS-receptors from *Vibrio* and *Legionella* are homodimers linked to histidine-kinases as cytosolic effector domains (***Ng and Bassler, 2009***). This is comparable to bacterial class III ACs which are homodimers (***Linder and Schultz, 2003***). These superficial observations suggested that by swapping the membrane anchor/receptor between a 6TM AC and CqsS from *Vibrio* one might generate a 6TM AC which is regulated by the QS-ligand CAI-1. We chose the mycobacterial class IIIa AC Rv1625c for this investigation because of its similarity to the mammalian congeners (***Guo et al., 2001***) and the QS-receptor from *V. harveyi*.

The success of generating a functionally productive chimera between the AC and the QS-receptor hinges on the precondition that a suitable point of transition between both proteins can be found which allows signal propagation from the CqsS receptor to the AC effector. The cytosolic aa sequences exiting from the respective last TMs have no conspicuous complementarity which would indicate a self-evident point of connection (***Figure 1***). Therefore, 15 different points of connection between the CqsS receptor and the catalytic domain of Rv1625c were probed. The first points of transition tested were the arginine residues present at the cytosolic membrane exit in both proteins (***Figure 1***). The chimera $CqsS_{1\text{-}168}Rv1625c_{203\text{-}443}$ was active (16.2 nmol cAMP·mg^{-1}·min^{-1}), but unregulated. The basal AC activity of this chimera was comparable to that of the membrane-bound Rv1625c holoenzyme (***Guo et al., 2001***), i.e. generally the two disparate domains were functionally fully compatible with each other. In the cyanobacterial class IIIa AC CyaG from *Arthrospira platensis* a distinct N-terminal domain that starts with RSEELL, was required for a functional interaction with the chemotaxis receptor Tsr (***Winkler et al., 2012***). A comparison between CyaG and Rv1625c AC sequences revealed a similar domain in Rv1625c beginning with RSEALL (***Figure 1***). Hence, for the Rv1625c AC this point of transition to CqsS was chosen. In the CqsS His-kinase the auto-phosphorylated histidine is part of the canonical H-box (underlined in ***Figure 1***; ***Grebe and Stock, 1999***). However, several chimeras of CqsS and Rv1625c linked in this region were unaffected by CAI-1. With transition points closer to the membrane exit of the CqsS receptor, e.g. at Val172, Ala181 or Gly185, AC activities were reproducibly stimulated by CAI-1 (not shown). For further experiments we linked Ala181 of the QS-receptor to Arg218 of the Rv1625c AC, generating $CqsS_{1\text{-}181}$-$Rv1625c_{218\text{-}443}$ (abbreviated CqsS-Rv1625c; ***Figure 1***) because it responded maximally. The chimera CqsS-Rv1625c was stimulated by 85% with 10 µM CAI-1 (***Figure 2***). The response was concentration-dependent and the EC_{50}

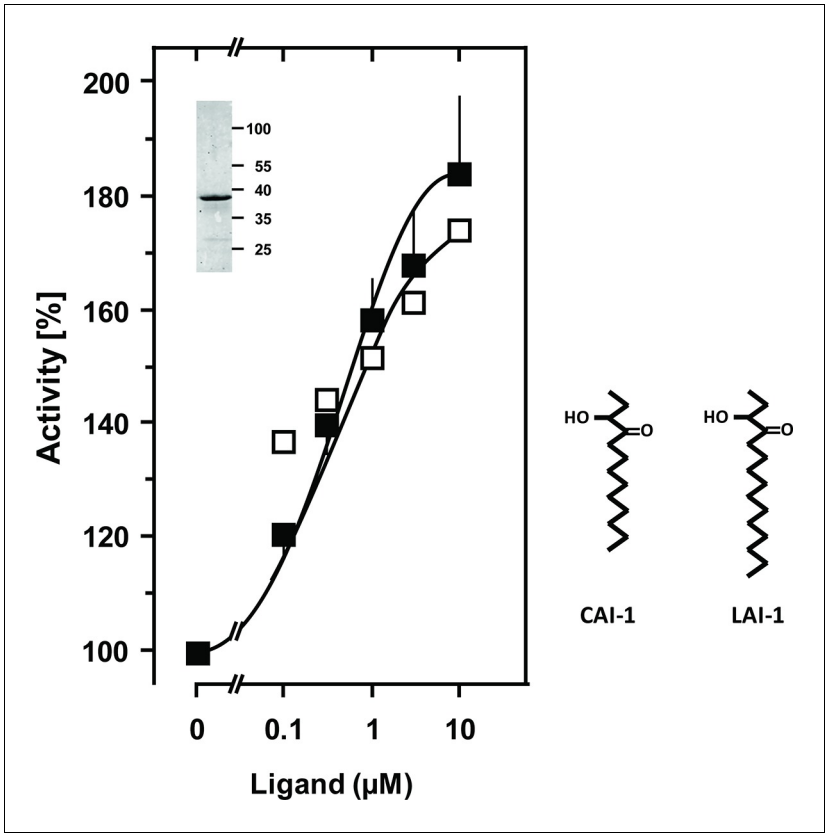


Figure 2. Stimulation of the chimera $CqsS_{1\text{-}181}$-$Rv1625c_{218\text{-}443}$ by the QS-ligands CAI-1 or LAI-1. Basal activity was 5.5 nmol cAMP·mg^{-1}·min^{-1}. The EC_{50} concentrations were 400 nM. Filled squares, CAI-1 (n = 5–12; ± S.E.M.); open squares, LAI-1 (n= 1–2). CAI-1 stimulations were significant starting at 100 nM ligand. The insert shows a Western blot of the expression product with MW standards indicated at the side. The structure of the ligands is depicted at right. The catalytic domain of Rv1625c alone was not affected by CAI-1 or LAI-1 (not shown).
DOI: 10.7554/eLife.13098.004

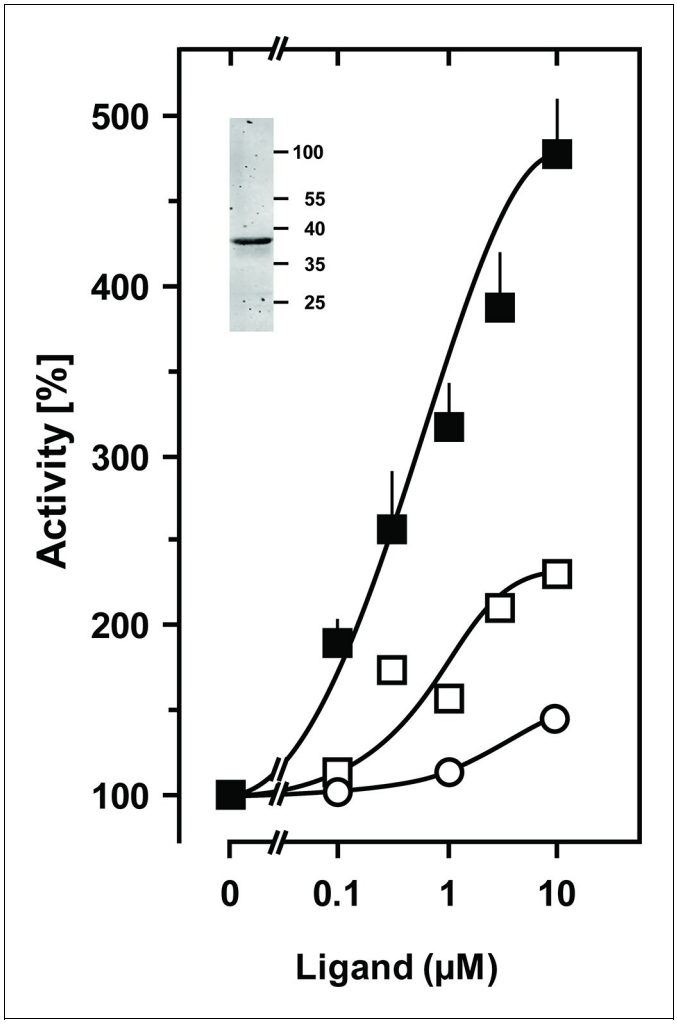


Figure 3. Stimulation of the chimera $CqsS_{1-181}$F166L-$Rv1625c_{218-443}$ by the QS-ligands CAI-1 or LAI-1. Basal activity was 4 nmol cAMP·mg^{-1}·min^{-1}. Filled squares, CAI-1 (n = 5–11; ± S.E.M.); open squares, LAI-1 (n=2); open circles, 3,4-tridecanediol. The EC_{50} concentrations were 400 nM CAI-1, 900 nM LAI-1, and 2000 nM 3,4-tridecanediol. CAI-1 stimulations were significant starting at 100 nM ligand. Insert: Western blot of expression product.
DOI: 10.7554/eLife.13098.005

concentration for CAI-1 was 400 nM (***Figure 2***). Discrimination between CAI-1 with a C9 and LAI-1 with a C11 lipid tail was absent. In both cases maximal stimulation at 10 µM CAI-1 and the EC_{50} concentrations were identical (***Figure 2***). These parameters perfectly matched the concentrations of 400 nM required to observe phenotypic responses from the CqsS-His-kinase in *V. cholerae* (***Ng et al., 2011***).

In vivo characterization of CqsS from *V. cholerae* identified Cys170 at the exit of TM6 as critical for signaling (***Ng et al., 2011***). In CqsS from *V. harveyi* the corresponding residue is Phe166 (***Figure 1***). We examined its role by substituting it with all 19 proteinogenic amino acids. Substitutions by large flexible hydrophobic residues, i.e. Leu, Ile, or Met enhanced CAI-1 stimulation up to 480% compared to 185% with the parent Phe166 construct (compare ***Figures 2*** and ***3***).

Furthermore, a high discrimination between CAI-1 and LAI-1 was effectuated and the potency ratio CAI-1/LAI-1 at 10 µM increased from 1 to 3 (see ***Figures 2*** and ***3***). Evidently, the F166L CqsS sensor domain in the chimera shows high selectivity between CAI-1 and LAI-1. Our findings were in line with earlier results that the natural CqsS receptor preferably detects ligands with C_{10} or C_8 but not shorter (C_6, C_4) or longer (C_{12}, e.g. LAI-1) alkyl tails (***Tiaden and Hilbi, 2012***). 3, 4-tridecanediol which has not been reported as a natural ligand stimulated by about 10% of CAI-1 at 10 µM (***Figure 3***). Other amino acid substitutions at position 166 of CqsS did not affect the extent of AC

stimulation with the exception of tryptophan which abrogated stimulation (data not shown). Activation enhanced Vmax from 14 to 42 nmol cAMP·mg^{-1}·min^{-1} whereas Km for substrate ATP was not significantly affected (233 and 163 µM ATP, respectively). An up to 10^4-fold activation of CqsS-stimulated reporter gene transcription was earlier observed in *V. harveyi* and *V. cholerae* (21). The 5-fold activation of CqsS-F166L reported here appears comparatively small. This discrepancy can be explained by the fact that the amplification systems used differ profoundly. Earlier studies investigated the quorum-sensing system *in vivo* with a reporter gene transcription/bioluminescence readout which is much more sensitive than an *in vitro* AC assay with isolated cell membranes used in the present study (**Ng et al., 2011**). In addition, the coupling of CqsS to its authentic effector might well be more stringent than that attainable in a chimera with an exogenous Rv1625c AC output domain.

Next we examined whether the CqsS-Rv1625c AC chimera is operational *in vivo*. Use of maltose by *E. coli* requires activation of the maltose operon via the cAMP/CRP signaling system. Maltose fermentation produces organic acids which are visualized on MacConkey plates by the pH indicator phenol red. We used the AC-deficient *E. coli cya-99* strain with a high affinity CRP variant (**Garges and Adhya, 1985**). It cannot metabolize carbohydrates for lack of cAMP. When grown on MacConkey agar, colonies appear whitish. We transformed the CqsS-F166L-Rv1625c into *E. coli cya-99*, plated the cells on MacConkey maltose agar and induced AC expression by a filter strip soaked with 30 mM IPTG. The reddish zone along the filter strip which is indicative of maltose metabolism was expanded at the side where 10 µl of a 100 µM CAI-1 solution was applied (**Figure 4**), clearly demonstrating a CAI-1 stimulated cAMP production *in vivo*.

CqsS dimerization is required for adenylate cyclase regulation

The monomeric bacterial class III ACs require homodimerization (**Linder and Schultz, 2003**). Similarly, bacterial His-kinases of two-component systems are homodimers in which a His-residue of the H-box is phosphorylated either in cis or trans (**Casino et al., 2014**; **2009**). Here we examined whether dimerization of the QS-receptors is required for AC regulation. The CqsS receptor was connected to known inactive Rv1625c AC point mutants, Rv1625cD300A and Rv1625cR376A. These point mutants complement each other and the dimer is catalytically active (**Guo et al., 2001**). Thus, CqsS-Rv1625cD300A was inserted into pETDuet-3 alone or in combination with CqsS-Rv1625cR376A. When CqsS-Rv1625cD300A and CqsS-Rv1625cR376A were jointly expressed, robust

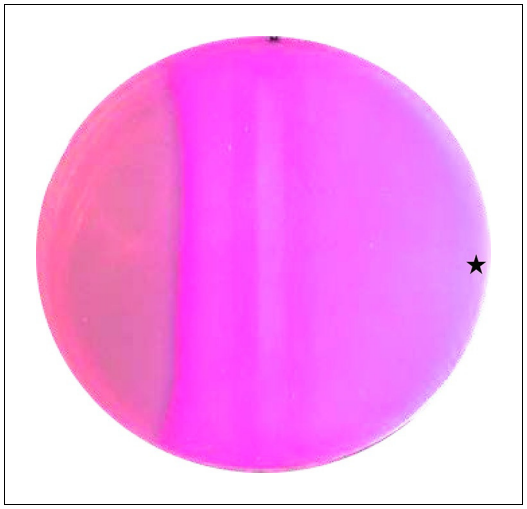


Figure 4. CAI-1 stimulates cAMP formation *in vivo*. A MacConkey maltose agar plate with *E. coli cya-99 crp**144 transformed with $CqsS_{1\text{-}181}$F166L-$Rv1625c_{218\text{-}443}$ was induced by a filter strip soaked with 1 mM IPTG (running from top to bottom in the middle). 10 µl of 100 µM CAI-1 in DMSO/water was spotted at the asterisk, the plate was tipped and the solution was allowed to move left. As a surface active compound it regularly spread over a large area. Note that the bacterial lawn at left was not induced. Picture was taken from the bottom of the Petri-dish (three independent experiments were carried out, each with at least three agar plates and different concentrations of CAI-1 and IPTG; controls with solvent were negative).
DOI: 10.7554/eLife.13098.006

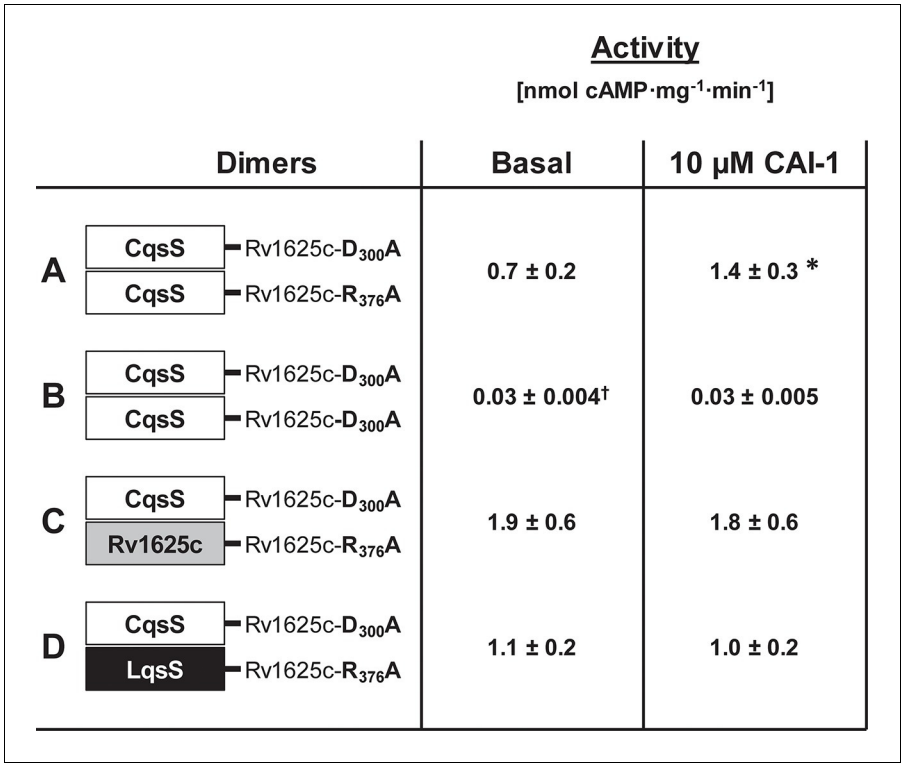


Figure 5. Homodimerization of the CqsS receptor is required for signaling. (**A**) with complementary Rv1625c point mutations Rv1626cD300A and Rv1625cR376A a regulated dimeric chimera was generated (*p<0.05 compared to respective basal activity). (**B**) as a control the construct CqsS-Rv1625cD300A was expressed alone. It was inactive. (**C**, **D**) complementing mutants with differing membrane domains were active, yet unregulated. Basal activity of construct B significantly differed from those in constructs **A**, **C**, and **D** (†p<0.001).
DOI: 10.7554/eLife.13098.007

AC activity and regulation by CAI-1 was observed (***Figure 5A***). This demonstrated that the membrane anchors of CqsS dimerize and allow individually inactive AC domains to form an active, ligand-regulated dimer. Expectedly, the expression product of CqsS-Rv1625cD300A was inactive (***Figure 5B***). The experiment did not answer the question whether ligand binding requires a CqsS dimer. This was addressed by linking the inactive Rv1625c monomers to the 6TM anchors of either Rv1625c or the *Legionella* LqsS QS-receptor. First, CqsS-Rv1625cD300A and the full-length Rv1625cR376A were jointly expressed in pETDuet-3. AC activity was observed, yet CAI-1 did not regulate (***Figure 5C***). Second, the QS-receptor from LqsS was joined with the inactive Rv1625cR376A in a similar manner. CqsS-Rv1625cD300A and LqsS-Rv1625cR376A were concomitantly expressed. As before, AC activity was restored, yet regulation by CAI-1 was absent (***Figure 5D***). This indicated that the membrane anchors were close enough to enable productive heterodimerization; however, a regulatory ligand-binding site was absent. Possibly one ligand molecule binds at the interface of a receptor homodimer as is the case in the chemotaxis receptors Tsr or Tar or Ni^{2+}-binding in PhoQ (***Cheung et al., 2008***; ***Gardina and Manson, 1996***; ***Kanchan et al., 2010***; ***Mowbray and Koshland, 1990***; ***Yang et al., 1993***). Because the highly lipophilic ligand does not allow meaningful receptor binding studies, presently this cannot be examined any further.

Stimulation of AC Rv1625c activity by CAI-1 is irreversible

The QS-system of *Vibrio* controls virulence. At high cell density, CAI-1 is produced, released and binds to the extracellular CqsS receptor (***Wei et al., 2012***). In a multistep intracellular process this results in reduced production of virulence factors and allows the pathogen to escape from the host, thus spreading disease (***Higgins et al., 2007***; ***Rutherford and Bassler, 2012***). Because extracellular loops for ligand binding are absent in CqsS and because of the lipophilicity of the ligand, CAI-1 may bind within the membrane segments of the receptor dimer. Therefore, the question whether CAI-1

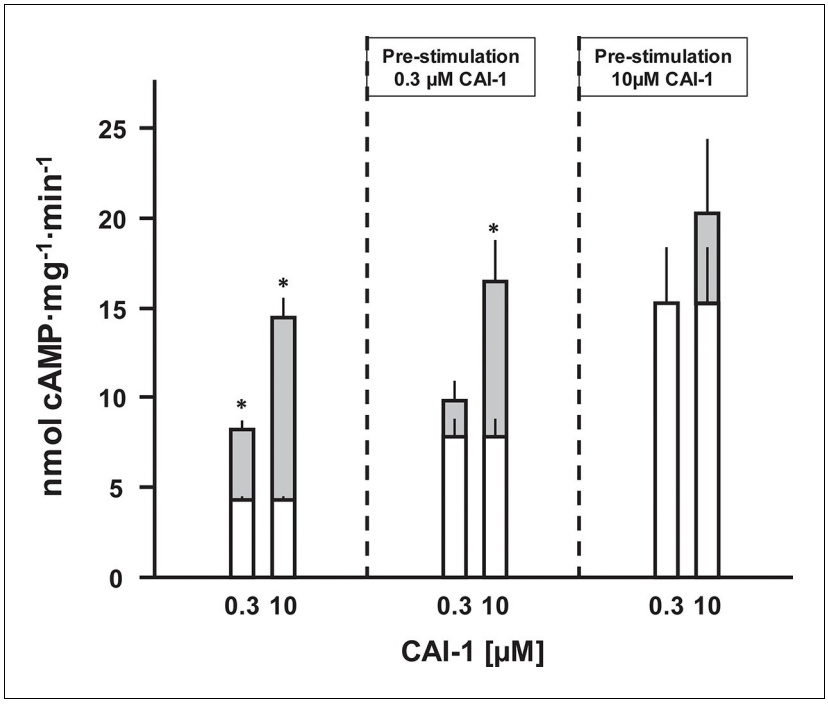


Figure 6. CAI-1 ligand binding to the CqsS QS-receptor is irreversible. Membranes containing $CqsS_{1\text{-}181}$F166L-$Rv1625c_{218\text{-}443}$ were stimulated with 0.3 or 10 µM CAI-1 (left), re-isolated and re-stimulated with 0.3 and 10 µM CAI-1. Only the stimulations marked with an asterisk differed significantly from the respective unstimulated controls. White bars represent basal AC activities, gray bars on top represent additive CAI-1-stimulated activities (S.E.M., n = 6).

DOI: 10.7554/eLife.13098.008

stimulation is reversible was examined next. Membranes were stimulated with sub-saturating (300 nM) and saturating (10 µM) concentrations of CAI-1 for 10 min. The reactions were rapidly stopped by cooling to 0˚C and the pre-stimulated membranes were re-isolated by ultracentrifugation. Those membranes were then stimulated again by CAI-1 (***Figure 6***). Membranes pre-treated with 300 nM CAI-1 had an elevated 'basal' AC activity equivalent to the previous 0.3 µM CAI-1 stimulation (***Figure 6***, left). Accordingly, re-stimulation by 0.3 µM CAI-1 failed whereas addition of 10 µM CAI-1 activated to the maximal possible extent. Membranes which were initially exposed to 10 µM CAI-1 remained almost fully activated and were completely refractory to re-stimulation (***Figure 6***). In this context we checked whether ligand was specifically binding at the receptor sites or remained unspecifically associated with the membranes. The supernatants of the ultracentrifugation steps of the pre-incubated membranes stimulated naïve membranes according to the previously tested CAI-1 concentrations. This excluded unspecific association of the lipophilic ligand with the membrane or adherence to the reaction vessels. Therefore, we can conclude that CAI-1 stimulation was irreversible. This is slightly reminiscent on the biochemistry of rhodopsin in the mammalian eye. There, retinal is even covalently bound into the membrane-segments of opsin, a GPCR with scant extramembrane loops. After excitation receptor regeneration requires ligand removal by an enzymatic process and transport to the retinal pigment epithelium which, notably, contains exclusively adenylate cyclase type VII (***Völkel et al., 1996***). How in *Vibrio* signal termination is accomplished remains to be investigated. Possibilities are an inactivating metabolism of CAI-1 or proteolysis.

6TM membrane anchors in ACs, universal receptor modules?

The above data demonstrated that the homodimeric catalytic domain of the canonical class IIIa Rv1625c AC was capable to decode the ligand-initiated conformational signal of a 6TM QS-receptor and translate it into a change in AC activity. The functional membrane anchor replacement accompanied by a gain of a novel physiological function suggests that the Rv1625c 6TM anchor actually constitutes an orphan receptor, i.e. receptor without known ligand. Therefore, we examined the

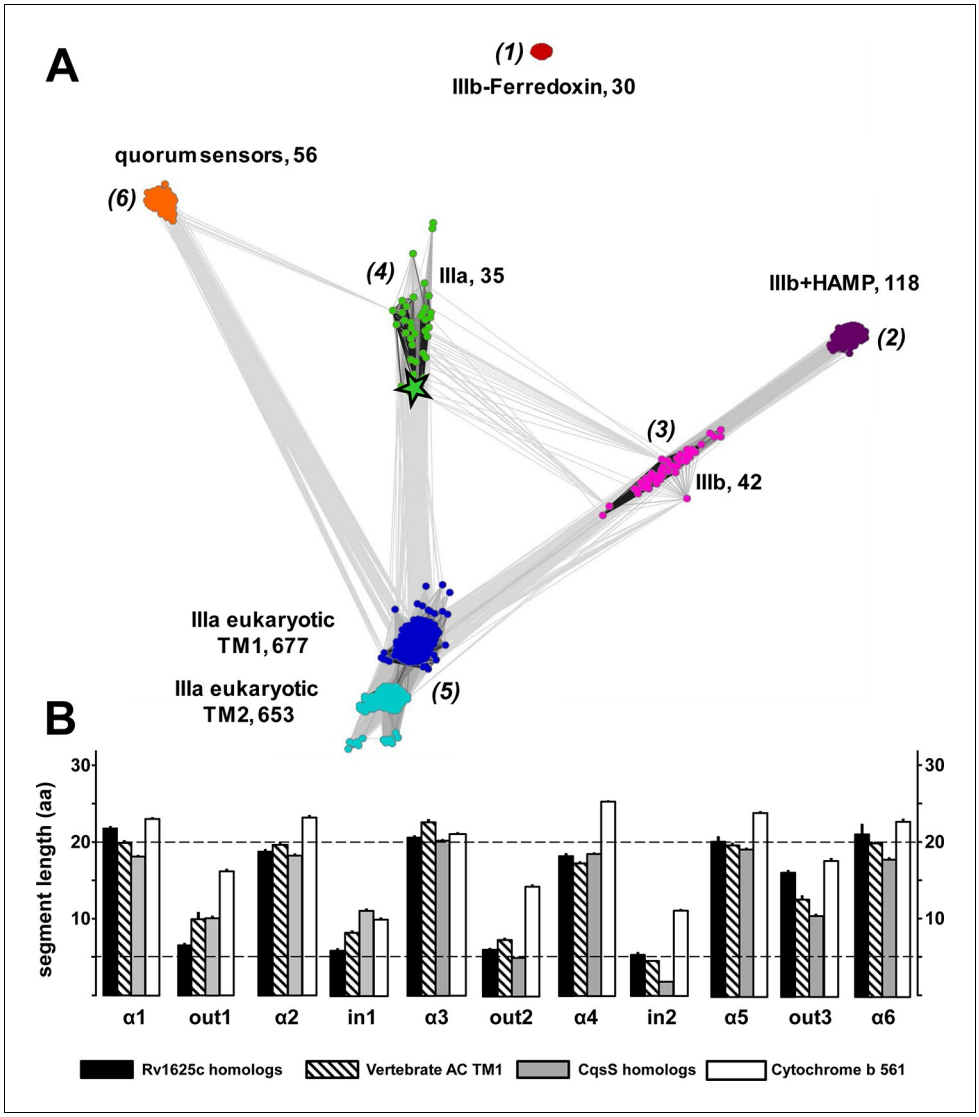


Figure 7. (**A**) Cluster map of 6TM domains of adenylate cyclases and CqsS-like sensory receptors. A comprehensive set of 6TM AC anchors was extracted from 408 eukaryotic and 1456 bacterial proteomes and clustered in CLANS (HHsearch p-value cutoff 5E-4, attraction value 10, repulsion value 5); outliers were removed. Each dot represents a single 6TM domain. Above threshold HHsearch hits are shown as connecting lines between AC pairs in different clusters; the darker line color, the more similar the protein sequences. 6TM anchors of ACs form five clusters of high pairwise sequence similarity: cluster (*1*), anchors of bacterial class IIIb ACs characterized by the presence of a cytosolic ferredoxin domain (30 sequences from the α and *β* branches of proteobacteria). Cluster (*2*), anchors of bacterial class IIIb ACs characterized by a signal-transducing HAMP domain (118 sequences mainly from Actinobacteria, but also from α-proteobacteria, δ-proteobacteria, Chlorobia and Thermoleophilia). Cluster (*3*), anchors of bacterial class IIIb ACs similar to HAMP-associated anchor domains but which lack a HAMP domain (42 sequences mainly from α-proteobacteria). Cluster (*4*), anchors from bacterial class IIIa ACs prototypically represented by the mycobacterial AC Rv1625c (35 from many different phyla of bacteria, including Actinobacteria, Proteobacteria, Chlorophyta, Spirochaetes, and Bacteriodetes). The enlarged asterisk denotes the position of the mycobacterial AC Rv1625c. Cluster (*5*), anchors of the pseudoheterodimeric eukaryotic class IIIa ACs (TM1 677 sequences, TM2 653 sequences). CqsS Cluster (*6*), 6TM domains of sensory His-Kinases similar to CqsS (from Bacteriodetes, Chlorobia, α-, *β*-, and γ-proteobacteria). (**B**) Length comparisons of the transmembrane helices and loops of 6TM membrane anchors/receptors/sensors. The data sets from clusters 4, 5 and 6 from ***Figure 7A*** were used supplemented with 250 cytochrome b561 proteins. In case no S.E.M. is visible the size of the vertical bar is within the line thickness of the respective bar. 'α'-Numbering denotes the consecutive TM helices starting from the N-terminus, 'out' and 'in' denote sequential extra- and intracellular loop sequences. The two horizontal lines are at the 5 and 20 aa level.

Figure 7 continued on next page

Figure 7 continued

DOI: 10.7554/eLife.13098.009

The following figure supplements are available for figure 7:

Figure supplement 1. Length comparisons of α-helices and loops of 6TM modules from adenylate cyclases and quorum sensors.

DOI: 10.7554/eLife.13098.010

Figure supplement 2. The 12 sequences used for the alignment were from *Agrobacterium albertimagni*, WP_006725538; *Arthrospira maxima*, B5VUZ0; *Arthrospira platensis*, D5A5G2; *Beggiatoa*, A7BXS6; *Dechloromonas*, Q47AI8; *Hyphomicrobium*, C6QBG1; *Lyngbya*, A0YQ82; *Mesorhizobium*, LSHC420B00; *Microcoleus sp.*, PCC_7113; *Oscillatoria acuminata*, PCC_6304; *Nostoc*, YP_001866931; *Mycobacterium tuberculosis*, ALB18789 (Rv1625c).

DOI: 10.7554/eLife.13098.011

Figure supplement 3. AC_1 Sequences used for the alignment: *Macaca fasci*, XP_005551359.1; *Bos taurus*, NP_776654.1; *Anas platyrhynchos*, XP_005014606.1; *Danio rerio*, NP_001161822.1; *Mesocricetus auratus*, XP_005083033.1; *Pseudopodoces humilis*, XP_005525280.1; *Gallus gallus*, XP_418883.4; *Homo sapiens*, NP_066939.1; *Mouse*, GI:62512159; *Heterocephalus glaber*, XP_004840600.1; *Orcinus orca*, XP_004283063.1; *Pan troglodytes*, XP_519081.3; *Ficedula albicollis*, XP_005041477.1; *Sarcophilus harrisii*, XP_003762599.1; *Odobenus rosmarus divergens*, XP_004409496.1; *Melopsittacus undulatus*, XP_005148091.1.

DOI: 10.7554/eLife.13098.012

relationship between the QS-receptors and the 6TM AC membrane anchors by a focused bioinformatic approach.

From the Uniprot Reference Proteomes databank we obtained all class III ACs using the highly conserved catalytic domain for data mining (see Materials and methods). The catalytic domains were stripped; the putative 6TM domains were extracted and retained. Similarly, we extracted 6TM modules related to the CqsS quorum-sensing receptor. The combined data set comprising a total of 1616 6TM modules was subjected to a cluster analysis using the CLANS software (***Figure 7A***; ***Frickey and Lupas, 2004***). In total, six distinct clusters were observed, five of which were interconnected to different extents, among them the cluster of QS-receptors (***Figure 7A***). This clear-cut separation of 6TM modules is surprising and revealing. Transmembrane spanning α-helices have a strong predilection for hydrophobic residues with Leu, Ile, Val, Ala, Gly and Phe comprising two-thirds (***Senes et al., 2000***). Yet, the unequivocal clustering indicates a differentiated pattern of conserved 6TM variants which is at odds with an assignment of simply an anchoring device. We propose that the diversity of conserved patterned subtypes mirrors a succinct adaptation to particular physiological functions. By far the majority of 6TM modules were eukaryotic ACs, from the slime mold *Dictyostelium*, the extant coelacanth up to man. The preponderance of eukaryotic sequences is, at least in part, due to the fact that in vertebrates nine distinct membrane-delimited pseudoheterodimeric AC isoforms exist, each with two TM modules, TM1 and TM2. The observed clusters were identified as follows: (*1*) An isolated, unconnected cluster of 30 bacterial class IIIb ACs is characterized by a ferredoxin between the membrane and the catalytic domain. A representative example is an AC from *Rhodopseudomonas palustris* (Uniprot Q132R4) which intracellularly appends to a ferredoxin module. The membrane anchors show sequence similarity to the cytochrome b subunits present in the fumarate reductase/succinate dehydrogenase protein families to an extent that they can be matched by simple BLAST searches. These 6TM modules have four precisely spaced intra-membrane histidine residues which coordinate heme as an electron carrier (***Einsle et al., 2000***; ***Hederstedt, 1998***; ***Kern et al., 2010a***; ***2010b***). It can confidently be predicted that the AC 6TM anchors of this cluster will turn out to contain two heme entities as prosthetic groups. (*2*) A cluster of 118 bacterial class IIIb ACs is characterized by a signal-transducing HAMP domain between membrane anchor and catalytic domain such as the well-studied mycobacterial AC Rv3645 (***Hazelbauer et al., 2008***; ***Hulko et al., 2006***; ***Kanchan et al., 2010***; ***Mondéjar et al., 2012***). (*3*) A cluster of 42 bacterial class IIIb ACs shows partly high local similarity to ACs of cluster (*2*) as visualized by the number and intensity of the gray colored connecting lines between both clusters (***Figure 7A***). Yet, ACs in cluster (*3*) lack a HAMP domain. To our knowledge, none of these ACs has ever been interrogated experimentally. (*4*) A cluster of 35 membrane anchors from bacterial class IIIa ACs as prototypically represented by the mycobacterial AC Rv1625c used here. It is predominantly connected to the cluster with mammalian TM modules. (*5*) Two tightly connected groups are combined here into one cluster of 677 TM1

and 653 TM2 modules which are derived from eukaryotic, mostly vertebrate, pseudoheterodimeric class IIIa ACs. The difference in the number of TM1 and TM2 units is due to individual sequence divergences of the TM2 domains, database miss-annotations, and problems in the automated prediction of transmembrane spans. (*6*) A cluster of 56 6TM modules corresponds to quorum-sensing receptors of the CqsS type.

The cluster of the eukaryotic TM1 domain (from cluster *5*) was connected to the clusters of the bacterial type IIIa (*4*) and type IIIb (*3*) by a large number of pairwise matches, typically covering all six transmembrane spans. In contrast, the CqsS cluster had significantly fewer and only local matches, aligning quorum-sensing sequences specifically to transmembrane spans 1 and 2 of individual eukaryotic AC TM1 sequences (all from the bony fish Osteichtyes), and to TM 3, 4 and 5 of the bacterial type IIIa ACs, each at approximately 20–25% identity. Notably, the average sequence identity within the clusters themselves is approximately 40%. Thus, currently we can only speculate whether the observed similarities indicate remote homology at the limit of recognition or convergence due to identical structural and functional constraints.

Visual inspection of sequence properties of single TMs indicated a remarkable shortness of α-helices and respective interconnecting loops already apparent in the 2D presentations (***Figure 1***). Notably, the loops are by far shorter than those of comparable sensory proteins, such as mammalian GPCRs with seven TMs, bacterial chemotaxis receptors, His-kinases and ACs with 2 TMs or 4 TMs such as the chase domain (**c**yclase/**h**istidine kinases **a**ssociated **s**ensory **e**xtracellular). To support this notion we investigated the length parameters of the transmembrane segments and loops of AC and CqsS anchor types which were used in the cluster analysis, and compared them to 250 orthologues of cytochrome b561 (***Figure 7B*** and ***Figure 7—figure supplement 1***). The latter protein was selected because a high-resolution structure is available, i.e. they form a 2x6TM homodimer of short TM spans and loops analogous to those from Rv1625c and CqsS (***Lu et al., 2014***). The lengths of the six TM segments and of the five loops were well conserved between orthologues of the same anchor type and highly similar between anchors of the bacterial AC classes IIIa and IIIb, eukaryotic TM1, and, most surprisingly, CqsS-like QS-receptors (***Figure 7B***). On the other hand, the eukaryotic TM2 anchors, those of the ferredoxin-associated class IIIb, and anchors coupled to a cytosolic HAMP-domain (class IIIb ACs in cluster 2) had one or even several elongated loops or α-helices (***Figure 7*** and ***Figure 7—figure supplement 1***).

The short helices of around 20 residues in this 6TM module design must cross the 30 Å thick lipid bilayer almost orthogonally and the short connecting loops restrict the number of positional permutations foreshadowing a compact packing. Therefore, the structures of transmembrane domains of eukaryotic AC TM1, the bacterial class IIIa ACs such as Rv1625c, and QS-receptors of the CqsS type can be reasoned to possess overall structural similarities with that of cytochrome b561.

Discussion

In 1989 the analysis of the first aa sequence of a mammalian AC suggested that the two membrane anchors, each consisting of six transmembrane segments, might carry a transporter or channel function (***Krupinski et al., 1989***). Since then, our knowledge concerning the membrane anchors has not advanced; to date no function for approximately 40% of a class IIIa AC protein sequence has been identified which goes beyond a mere membrane fixation. After 26 years we demonstrate for the first time that a canonical class IIIa AC with a 6TM membrane anchor is directly regulated by a membrane receptor of an identical design, yet with a known ligand, the QS-receptor from *V. harveyi*.

Our data raise novel questions concerning the evolution of the regulation of class III ACs with 6TM membrane-anchoring modules. In absence of bacterial G-proteins bacterial class III ACs probably will turn out to be directly regulated via their prominent membrane anchors with the above mentioned *R. palustris* AC as an emerging example. In fact the membrane anchors of the class IIIa bacterial ACs are highly diverged whereas the catalytic domains are conserved (***Figure 7*** and ***Figure 7—figure supplement 2***). This suggests that in bacteria the TM domains have evolved independently and very rapidly relative to the catalytic domains. This is in agreement with the general observation that mutations in upstream regulatory domains mostly are neutral and over time a variety of ligand specificities have evolved by chance mutations. At the same time even slightly detrimental mutations in downstream effector domains are not tolerated, thus resulting in sequence and functional conservation (***Schultz and Natarajan, 2013***). In such an evolutionary scenario the

mechanisms of signal transduction per se remain intact and allow revealing combinatorial diversity as demonstrated here (**Schultz et al., 2015**).

In contrast, in vertebrates the membrane-delimited ACs are regulated indirectly by GPCR activation which intracellularly results in release of Gα proteins. Is this the result of a loss of direct ligand regulation during evolution while indirect GPCR regulation evolved or has ligand regulation of mammalian ACs been missed so far? Because the AC membrane bundles of vertebrates are highly conserved in an isoform-specific manner from the coelacanth to man (**Figure 7** and **Figure 7—figure supplement 3** as an example of mammalian AC_I and AC_II subtypes) one can reasonably assume that they have an indispensable physiological function. This could be a particular compartmentalized membrane localization e.g. described in (**Crossthwaite et al., 2005**) or, in our view more likely and in accordance with our findings an as yet hidden receptor function, as a direct ligand-binding module or in conjunction with an accessory membrane protein that operates as the true sensor. The second alternative would suggest that in vertebrates regulation of intracellular cAMP concentrations is subject to an interaction between direct, ligand-mediated and indirect GPCR-Gsα-regulated effects. A lasting receptor occupation by a ligand during different physiological states might well set the responsiveness of mammalian AC isoforms to a transient GPCR-Gsα activation. This would allow that lasting and transient physiological conditions converge via direct and indirect regulation in a central second messenger system, a situation absent in bacteria. This concept poses the questions whether in vertebrates suitable extracellular signals exist and whether the molecular provisions for a direct signal transduction through the 9 AC isoforms have been evolutionarily conserved. We are currently exploring these questions.

Recently we demonstrated that we could functionally replace the bacterial AC 6TM anchors by the *E. coli* chemotaxis receptors for serine or aspartate (**Kanchan et al., 2010**; **Schultz et al., 2015**; **Winkler et al., 2012**). By analogy, the results supported the hypothesis of a receptor function for the AC membrane anchors. A less plausible interpretation would have been that the functional coupling might just be a manifestation of the modular composition of signaling proteins. The compatibility may only have required that a satisfactory domain order is maintained in such chimeras without invoking a functional relatedness. The data reported here add an entirely novel dimension to the working hypothesis that AC membrane anchors in bacterial homodimeric as well as in mammalian pseudoheterodimeric ACs function as ligand receptors. The 2D models of the membrane anchors of Rv1625c and CqsS are similar (see **Figure 1**). This should allow an almost isosteric replacement. Competent QS-receptor and cyclase chimeras were dependent on the point of linkage indicating that the connecting sites had been evolutionarily predesigned, were conserved and fully operational for signal transduction between functionally differing proteins (**Schultz and Natarajan, 2013**). Implicitly this supports the prediction that the 6TM anchors of such ACs have a receptor function for which stimuli have yet to be identified. Because the ligands for CqsS and LqsS are known, the direct regulation of AC activity then is no real surprise. Admittedly, based alone on the biochemical data one might again take pains to argue that the exchangeability of 6TM-anchors just extends the range of signaling modules which can structurally and possibly functionally replace each other. This alternative interpretation is rather implausible in our view. First mixing and matching TM domains of eukaryotic AC is impossible without loss of activity (**Seebacher et al., 2001**). Second, the cluster analysis visualizes similarities between individual pairs of CqsS receptor type modules and membrane anchors from class IIIa ACs. Thus it supports the hypothesis that the membrane anchors of class III ACs, bacterial and mammalian alike have a function beyond membrane fixation. In this context cluster (*1*), which is unrelated to all other clusters, is particularly interesting. It demonstrates that comparable transmembrane architectures can result in significantly different sequence patterns notwithstanding the similar amino acid composition commonly shared by all membrane domains. Currently, it is impossible to speculate about or even predict the nature of ligands for the 6TM AC modules examined here. One might expect, however, that they will be closely associated with the intra-membrane space as extra-membranous loops for ligand-binding are noticeably absent in this type of 6TM bundles.

Materials and methods

An *E.coli* culture containing the protein CqsS of *V. harveyi* was obtained from K. Jung and LqsS from H. Hilbi, LMU, Munich (Germany). *M. tuberculosis* AC Rv1625c and point mutants D300A and R376A

were available in the laboratory (**Guo et al., 2001**). Radiochemicals were from Hartmann Analytik (Braunschweig, Germany) and Perkin Elmer (Rodgau, Germany). Enzymes were from either New England Biolabs or Roche Diagnostics. Other chemicals were from Sigma, Roche Diagnostics, Merck and Roth. CAI-1, LAI-1 and 3,4-tridecanediol were synthesized in-house according to (**Bolitho et al., 2011**; **Ng et al., 2010**). Concentration-response curves were usually limited to maximally 10 µM as the ligands have surfactant properties and assays with higher concentrations tended to give inconsistent results.

Plasmid construction

In CqsS-Rv1625c chimeras the following CqsS receptor length variants were probed in different combinations: F168; V172; K177; A181; S183; G185; G187; I188; H190; P195;L196. For AC Rv1625c catalytic domains the tested length variants were: A201, L202, R203 and R218. Standard molecular biology methods were used for DNA manipulations (primer sequences are in **Supplementary file 1**). DNA fragments and vectors were restricted at their 5´BamHI or EcoRI and 3´HindIII sites and inserted into pQE80$_L$ (Δ XhoI; Δ NcoI). When appropriate, silent restriction sites were introduced. All constructs carried an N-terminal His$_6$-tag for detection in Western blots. In the pETDuet-3 vector the first MCS was been replaced by that of pQE30 introducing an N-terminal His$_6$-tag. The second MCS in pETDuet-3 carried a C-terminal S-tag for Western blotting. The fidelity of all constructs was confirmed by double-stranded DNA sequencing.

Protein expression

Constructs were transformed into *E. coli* BL21(DE3). Strains were grown overnight in LB medium (20g LB broth/l) at 37˚C containing 100 µg/ml ampicillin. 200 ml LB medium (with antibiotic) was inoculated with 5 ml of a preculture and grown at 37˚C. At an A_{600} of 0.3, the temperature was lowered to 22˚C and the expression was started with 500 µM isopropyl thio-β-D-galactoside (IPTG) for 2.5–5 hrs. Cells were harvested by centrifugation, washed once with 50 mM Tris/HCl, 1mM EDTA, pH 8 and stored at -80˚C. For preparation of cell membranes cells were suspended in lysis buffer (50 mM Tris/HCl, 2 mM 3-thioglycerol, 50 mM NaCl, pH 8) containing complete protease inhibitor cocktail (Roche Molecular, Mannheim, Germany) and disintegrated by a French press (1100 p.s.i.). After removal of cell debris (4.300 x g, 30min, 4˚C) membranes were collected at 100000 x g (1h at 4˚C). Membranes were suspended in buffer (40 mM Tris/HCl, pH8, 1.6 mM 3-thioglycerol, 20% glycerol) and assayed for AC activity. A more detailed description of protein expression and membrane preparation is available as described in detail at Bio-protocol (**Beltz and Schultz, 2016**).

Adenylate cyclase assay

Adenylyl cyclase activity was determined for 10 min in 100 µl at 37˚C (**Salomon et al., 1974**). The reactions contained 5 µg protein, 50 mM Tris/HCl pH 8, 22% glycerol, 3 mM $MnCl_2$, 6 mM creatine phosphate and 230 µg creatine kinase, 75 µM [α-^{32}P]-ATP, and 2 mM [2,8-^{3}H]-cAMP to monitor yield during cAMP purification. Substrate conversion was kept below 10%.

Western blot analysis

The integrity of expressed recombinant membrane proteins was probed by Western blotting. Sample buffer was added to the membrane fractions and applied to SDS-PAGE (12 or 15%), in which proteins were separated according to size. For Western blot analysis, proteins were blotted onto PVPF membrane and examined with an RGS-His$_4$-antibody (Qiagen, Hilden, Germany) or S-tag antibody (Novagen R&D systems, Darmstadt, Germany) and a 1:2500 dilution of the fluorophore conjugated secondary antibody Cy3 (ECL Plex goat-α-mouse IgG-Cy3, GE Healthcare, Freiburg, Germany). Detection was carried out with the Ettan DIGE Imager (GE Healthcare). In general, proteolysis of expressed proteins was not observed.

Bioinformatics

Dataset

Sequences were taken from the Uniprot Reference Proteomes databank (release 2015_04). AC anchor sequences were identified by their conserved catalytic domain in a HMMer3 search (E-value cutoff 1E-5; (**Eddy, 2011**) with EBI SMART's cycc family alignment (**Schultz et al., 1998**). To extract

putative membrane domains, all segments N-terminal of cycc hits, up to the protein N-terminus or another cycc domain, were extracted. HAMP domains, if present, were located using HMMer3 and the HAMP EBI SMART family alignment and then removed. The extracted sequences were clustered to 30% sequence identity using kClust (**Hauser et al., 2013**). Then, the clusters were aligned individually and their TM spans were predicted by Polyphobius (**Kall et al., 2005**) and manual inspection. 6TM clusters were merged and used for further analyses. CqsS and cytochrome b561 homologs were identified in HMMer3 jackhmmer searches, using the *V. cholerae* CqsS and *A. thaliana* Cy561 anchors (Uniprot identifiers Q9KM66 and Q9SWS1) as queries.

Cluster analysis

For every sequence in the 6TM AC anchor data set we searched the Uniprot database, clustered at 20% sequence identity (Uniprot20, June 2015, available at http://toolkit.tuebingen.mpg.de), with HHblits (p-value cutoff 1E-3, minimum coverage 50%, two iterations; (**Remmert et al., 2012**). The resulting sequence alignments were used to create profile hidden Markov models (HMMs) that included Polyphobius TM predictions instead of secondary structure annotation. The same procedure was applied to CqsS anchor homologs. The clustering was performed using the CLANS software (**Frickey and Lupas, 2004**), based on pairwise HMM-HMM comparisons (**Soding, 2005**).

Helix/loop length analysis

The helix and loop lengths were measured using the sequences taken from the respective clusters of the CLANS analysis and cytochrome b561 homologs. Predicted transmembrane spans were based on manually refined Polyphobius predictions.

Statistical analysis

All experiments were repeated at least thrice. Data are presented as means ± S.E.M. when applicable. Student's t test was used.

Acknowledgements

This publication is dedicated to Prof. Dr. Günter Schultz, FU Berlin, at the occasion of his 80th birthday. Supported by Deutsche Forschungsgemeinschaft (SFB 766; TP B08) and institutional funds of the Max-Planck-Society. We are grateful to Prof. Dr. K. Hantke who carried out the *in vivo* experiments and to Prof. Dr. P. Koch for the synthesis of 3, 4-tridecanediol. We thank A. Schultz for invaluable technical assistance. We are indebted to Prof. Dr. A. Lupas for continuous encouragement and critical discussions.

Additional information

Funding

Funder	Grant reference number	Author
Deutsche Forschungsgemeinschaft	SFB 766	Stephanie Beltz
Max-Planck-Gesellschaft		Jens Bassler

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

SB, JB, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; JES, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article

Author ORCIDs

Joachim E Schultz, http://orcid.org/0000-0002-1985-4853

Additional files

Supplementary files

• Supplementary file 1. List of sense (s) and antisense (as) primers used for generating the diversity of chimeras used in this study.
DOI: 10.7554/eLife.13098.013

References

Beltz S, Schultz JE. , 2016. Protein expression protocol for an adenylate cyclase anchored by a Vibrio quorum sensing receptor. *Bio-protocol* **7**:e2112. doi: 10.21769/BioProtoc.2112

Bolitho ME, Perez LJ, Koch MJ, Ng W-L, Bassler BL, Semmelhack MF. 2011. Small molecule probes of the receptor binding site in the Vibrio cholerae CAI-1 quorum sensing circuit. *Bioorganic & Medicinal Chemistry* **19**: 6906–6918. doi: 10.1016/j.bmc.2011.09.021, PMID: 22001326

Bârzu O, Danchin A. 1994. Adenylyl cyclases: a heterogeneous class of ATP-utilizing enzymes. *Progress in Nucleic Acid Research and Molecular Biology* **49**:241–283. PMID: 7863008

Casino P, Miguel-Romero L, Marina A. 2014. Visualizing autophosphorylation in histidine kinases. *Nature Communications* **5**:3258. doi: 10.1038/ncomms4258, PMID: 24500224

Casino P, Rubio V, Marina A. 2009. Structural Insight into Partner Specificity and Phosphoryl Transfer in Two-Component Signal Transduction. *Cell* **139**:325–336. doi: 10.1016/j.cell.2009.08.032, PMID: 19800110

Chang G, Roth CB. 2001. Structure of MsbA from E. coli: A Homolog of the Multidrug Resistance ATP Binding Cassette (ABC) Transporters. *Science* **293**:1793–1800. doi: 10.1126/science.293.5536.1793, PMID: 11546864

Cheung J, Bingman CA, Reyngold M, Hendrickson WA, Waldburger CD. 2008. Crystal Structure of a Functional Dimer of the PhoQ Sensor Domain. *Journal of Biological Chemistry* **283**:13762–13770. doi: 10.1074/jbc.M710592200, PMID: 18348979

Crossthwaite AJ, Seebacher T, Masada N, Ciruela A, Dufraux K, Schultz JE, Cooper DMF. 2005. The Cytosolic Domains of Ca2+-sensitive Adenylyl Cyclases Dictate Their Targeting to Plasma Membrane Lipid Rafts. *Journal of Biological Chemistry* **280**:6380–6391. doi: 10.1074/jbc.M411987200, PMID: 15574428

Eddy SR. 2011. Accelerated Profile HMM Searches. *PLoS Computational Biology* **7**:e1002195. doi: 10.1371/journal.pcbi.1002195, PMID: 22039361

Einsle O, Stach P, Messerschmidt A, Simon J, Kroger A, Huber R, Kroneck PMH. 2000. Cytochrome c nitrite reductase from Wolinella succinogenes: Structure at 1.6 a Resolution, Inhibitor Binding, and Heme-Packing Motifs. *Journal of Biological Chemistry* **275**:39608–39616. doi: 10.1074/jbc.M006188200, PMID: 10984487

Frickey T, Lupas A. 2004. CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20**:3702–3704. doi: 10.1093/bioinformatics/bth444, PMID: 15284097

Gardina PJ, Manson MD. 1996. Attractant Signaling by an Aspartate Chemoreceptor Dimer with a Single Cytoplasmic Domain. *Science* **274**:425–426. doi: 10.1126/science.274.5286.425, PMID: 8832892

Garges S, Adhya S. 1985. Sites of allosteric shift in the structure of the cyclic AMP receptor protein. *Cell* **41**:745–751. doi: 10.1016/S0092-8674(85)80055-6, PMID: 2988785

Grebe TW, Stock JB. 1999. The histidine protein kinase superfamily. *Advances in Microbial Physiology* **41**:139–227. doi: 10.1016/S0065-2911(08)60167-8, PMID: 10500846

Guo YL, Seebacher T, Kurz U, Linder JU, Schultz JE. 2001. Adenylyl cyclase Rv1625c of Mycobacterium tuberculosis: a progenitor of mammalian adenylyl cyclases. *The EMBO Journal* **20**:3667–3675. doi: 10.1093/emboj/20.14.3667, PMID: 11447108

Hauser M, Mayer CE, Söding J. 2013. kClust: fast and sensitive clustering of large protein sequence databases. *BMC Bioinformatics* **14**:248–259. doi: 10.1186/1471-2105-14-248, PMID: 23945046

Hazelbauer GL, Falke JJ, Parkinson JS. 2008. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends in Biochemical Sciences* **33**:9–19. doi: 10.1016/j.tibs.2007.09.014, PMID: 18165013

Hederstedt L. 1998. The membrane-integral domain of succinate:quinone oxidoreductases –- a secretive haem-containing domain. *Biochemical Society Transactions* **26**:408–413. doi: 10.1042/bst0260408, PMID: 9765888

Higgins DA, Pomianek ME, Kraml CM, Taylor RK, Semmelhack MF, Bassler BL. 2007. The major Vibrio cholerae autoinducer and its role in virulence factor production. *Nature* **450**:883–886. doi: 10.1038/nature06284, PMID: 18004304

Hulko M, Berndt F, Gruber M, Linder JU, Truffault V, Schultz A, Martin J, Schultz JE, Lupas AN, Coles M. 2006. The HAMP Domain Structure Implies Helix Rotation in Transmembrane Signaling. *Cell* **126**:929–940. doi: 10.1016/j.cell.2006.06.058, PMID: 16959572

Kall L, Krogh A, Sonnhammer ELL. 2005. An HMM posterior decoder for sequence feature prediction that includes homology information. *Bioinformatics Bioinformatics* **2121**:i251251–i257-257. doi: 10.1093/bioinformatics/bti1014

Kanchan K, Linder J, Winkler K, Hantke K, Schultz A, Schultz JE. 2010. Transmembrane Signaling in Chimeras of the Escherichia coli Aspartate and Serine Chemotaxis Receptors and Bacterial Class III Adenylyl Cyclases. *Journal of Biological Chemistry* **285**:2090–2099. doi: 10.1074/jbc.M109.051698, PMID: 19923210

Kern M, Eisel F, Scheithauer J, Kranz RG, Simon J. 2010a. Substrate specificity of three cytochrome *c* haem lyase isoenzymes from *Wolinella succinogenes* : unconventional haem *c* binding motifs are not sufficient for haem *c*

attachment by NrfI and *Ccs*A1. *Molecular Microbiology* **75**:122–137. doi: 10.1111/j.1365-2958.2009.06965.x, PMID: 19919672

Kern M, Scheithauer J, Kranz RG, Simon J. 2010b. Essential histidine pairs indicate conserved haem binding in epsilonproteobacterial cytochrome c haem lyases. *Microbiology* **156**:3773–3781. doi: 10.1099/mic.0.042838-0, PMID: 20705660

Khandelwal RL, Hamilton IR. 1971. Purification and properties of adenyl cyclase from Streptococcus salivarius. *The Journal of Biological Chemistry* **246**:3297–3304. PMID: 4324896

Kleinboelting S, Diaz A, Moniot S, van den Heuvel J, Weyand M, Levin LR, Buck J, Steegborn C. 2014. Crystal structures of human soluble adenylyl cyclase reveal mechanisms of catalysis and of its activation through bicarbonate. *Proceedings of the National Academy of Sciences of the United States of America* **111**:3727–3732. doi: 10.1073/pnas.1322778111, PMID: 24567411

Kleven MD, Dlakić M, Lawrence CM. 2015. Characterization of a Single b-type Heme, FAD, and Metal Binding Sites in the Transmembrane Domain of Six-transmembrane Epithelial Antigen of the Prostate (STEAP) Family Proteins. *Journal of Biological Chemistry* **290**:22558–22569. doi: 10.1074/jbc.M115.664565, PMID: 26205815

Krupinski J, Coussen F, Bakalyar H, Tang W, Feinstein P, Orth K, Slaughter C, Reed R, Gilman A. 1989. Adenylyl cyclase amino acid sequence: possible channel- or transporter-like structure. *Science* **244**:1558–1564. doi: 10.1126/science.2472670, PMID: 2472670

Linder JU, Schultz JE. 2003. The class III adenylyl cyclases: multi-purpose signalling modules. *Cellular Signalling* **15**:1081–1089. doi: 10.1016/S0898-6568(03)00130-X, PMID: 14575863

Linder JU, Schultz JE. 2008. Versatility of signal transduction encoded in dimeric adenylyl cyclases. *Current Opinion in Structural Biology* **18**:667–672. doi: 10.1016/j.sbi.2008.11.008, PMID: 19054664

Lu P, Ma D, Yan C, Gong X, Du M, Shi Y. 2014. Structure and mechanism of a eukaryotic transmembrane ascorbate-dependent oxidoreductase. *Proceedings of the National Academy of Sciences of the United States of America* **111**:1813–1818. doi: 10.1073/pnas.1323931111, PMID: 24449903

Mates AK, Sayed AK, Foster JW. 2007. Products of the Escherichia coli Acid Fitness Island Attenuate Metabolite Stress at Extremely Low pH and Mediate a Cell Density-Dependent Acid Resistance. *Journal of Bacteriology* **189**:2759–2768. doi: 10.1128/JB.01490-06, PMID: 17259322

Mondéjar LG, Lupas A, Schultz A, Schultz JE. 2012. HAMP Domain-mediated Signal Transduction Probed with a Mycobacterial Adenylyl Cyclase as a Reporter. *Journal of Biological Chemistry* **287**:1022–1031. doi: 10.1074/jbc.M111.284067, PMID: 22094466

Mowbray SL, Koshland DE. 1990. Mutations in the aspartate receptor of Escherichia coli which affect aspartate binding. *The Journal of Biological Chemistry* **265**:15638–15643. PMID: 2168420

Ng W-L, Bassler BL. 2009. Bacterial Quorum-Sensing Network Architectures. *Annual Review of Genetics* **43**:197–222. doi: 10.1146/annurev-genet-102108-134304, PMID: 19686078

Ng W-L, Perez LJ, Wei Y, Kraml C, Semmelhack MF, Bassler BL. 2011. Signal production and detection specificity in Vibrio CqsA/CqsS quorum-sensing systems. *Molecular Microbiology* **79**:1407–1417. doi: 10.1111/j.1365-2958.2011.07548.x, PMID: 21219472

Ng W-L, Wei Y, Perez LJ, Cong J, Long T, Koch M, Semmelhack MF, Wingreen NS, Bassler BL. 2010. Probing bacterial transmembrane histidine kinase receptor-ligand interactions with natural and synthetic molecules. *Proceedings of the National Academy of Sciences of the United States of America* **107**:5575–5580. doi: 10.1073/pnas.1001392107, PMID: 20212168

Remmert M, Biegert A, Hauser A, Söding J. 2012. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nature Methods* **9**:173–175. doi: 10.1038/nmeth.1818

Rutherford ST, Bassler BL. 2012. Bacterial Quorum Sensing: Its Role in Virulence and Possibilities for Its Control. *Cold Spring Harbor Perspectives in Medicine* **2**:a012427. doi: 10.1101/cshperspect.a012427, PMID: 23125205

Salomon Y, Londos C, Rodbell M. 1974. A highly sensitive adenylate cyclase assay. *Analytical Biochemistry* **58**:541–548. doi: 10.1016/0003-2697(74)90222-X, PMID: 4827395

Schultz J, Milpetz F, Bork P, Ponting CP. 1998. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences of the United States of America* **95**:5857–5864. . doi: 10.1073/pnas.95.11.5857

Schultz JE, Kanchan K, Ziegler M. 2015. Intraprotein signal transduction by HAMP domains: A balancing act. *International Journal of Medical Microbiology* **305**:243–251. doi: 10.1016/j.ijmm.2014.12.007, PMID: 25595022

Schultz JE, Natarajan J. 2013. Regulated unfolding: a basic principle of intraprotein signaling in modular proteins. *Trends in Biochemical Sciences* **38**:538–545. doi: 10.1016/j.tibs.2013.08.005, PMID: 24051046

Seebacher T, Linder JU, Schultz JE. 2001. An isoform-specific interaction of the membrane anchors affects mammalian adenylyl cyclase type V activity. *European Journal of Biochemistry* **268**:105–110. doi: 10.1046/j.1432-1327.2001.01850.x, PMID: 11121109

Senes A, Gerstein M, Engelman DM. 2000. Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions. *Journal of Molecular Biology* **296**:921–936. doi: 10.1006/jmbi.1999.3488, PMID: 10677292

Soding J. 2005. Protein homology detection by HMM-HMM comparison. *Bioinformatics Bioinformatics* **2121**:951 951–960-960. doi: 10.1093/bioinformatics/bti125, PMID: 15531603

Spirig T, Tiaden A, Kiefer P, Buchrieser C, Vorholt JA, Hilbi H. 2008. The Legionella Autoinducer Synthase LqsA Produces an -Hydroxyketone Signaling Molecule. *Journal of Biological Chemistry* **283**:18113–18123. doi: 10.1074/jbc.M801929200, PMID: 18411263

Steegborn C, Litvin TN, Levin LR, Buck J, Wu H. 2005. Bicarbonate activation of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment. *Nature Structural & Molecular Biology* **12**:32–37. doi: 10.1038/nsmb880, PMID: 15619637

Sutherland EW, Rall TW. 1958. Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles. *The Journal of Biological Chemistry* **232**:1077–1091. PMID: 13549488

Tesmer JJ, Sunahara RK, Gilman AG, Sprang SR. 1997. Crystal Structure of the Catalytic Domains of Adenylyl Cyclase in a Complex with Gsalpha. *GTPgammaS. Science* **278**:1907–1916. doi: 10.1126/science.278.5345.1907, PMID: 9417641

Tews I, Findeisen F, Sinning I, Schultz A, Schultz JE, Linder JU. 2005. The Structure of a pH-Sensing Mycobacterial Adenylyl Cyclase Holoenzyme. *Science* **308**:1020–1023. doi: 10.1126/science.1107642, PMID: 15890882

Tiaden A, Hilbi H. 2012. α-Hydroxyketone Synthesis and Sensing by Legionella and Vibrio. *Sensors* **12**:2899–2919. doi: 10.3390/s120302899, PMID: 22736983

Völkel H, Beitz E, Klumpp S, Schultz JE. 1996. Cloning and expression of a bovine adenylyl cyclase type VII specific to the retinal pigment epithelium. *FEBS Letters* **378**:245–249. doi: 10.1016/0014-5793(95)01470-5, PMID: 8557110

Wei Y, Ng W-L, Cong J, Bassler BL. 2012. Ligand and antagonist driven regulation of the Vibrio cholerae quorum-sensing receptor CqsS. *Molecular Microbiology* **83**:1095–1108. doi: 10.1111/j.1365-2958.2012.07992.x, PMID: 22295878

Winkler K, Schultz A, Schultz JE. 2012. The S-Helix Determines the Signal in a Tsr Receptor/Adenylyl Cyclase Reporter. *Journal of Biological Chemistry* **287**:15479–15488. doi: 10.1074/jbc.M112.348409, PMID: 22427653

Yang Y, Park H, Inouye M. 1993. Ligand Binding Induces an Asymmetrical Transmembrane Signal through a Receptor Dimer. *Journal of Molecular Biology* **232**:493–498. doi: 10.1006/jmbi.1993.1405, PMID: 8393938

Yankovskaya V, Horsefield R, Törnroth S, Luna-Chavez C, Miyoshi H, Léger C, Byrne B, Cecchini G, Iwata S. 2003. Architecture of Succinate Dehydrogenase and Reactive Oxygen Species Generation. *Science* **299**:700–704. doi: 10.1126/science.1079605, PMID: 12560550

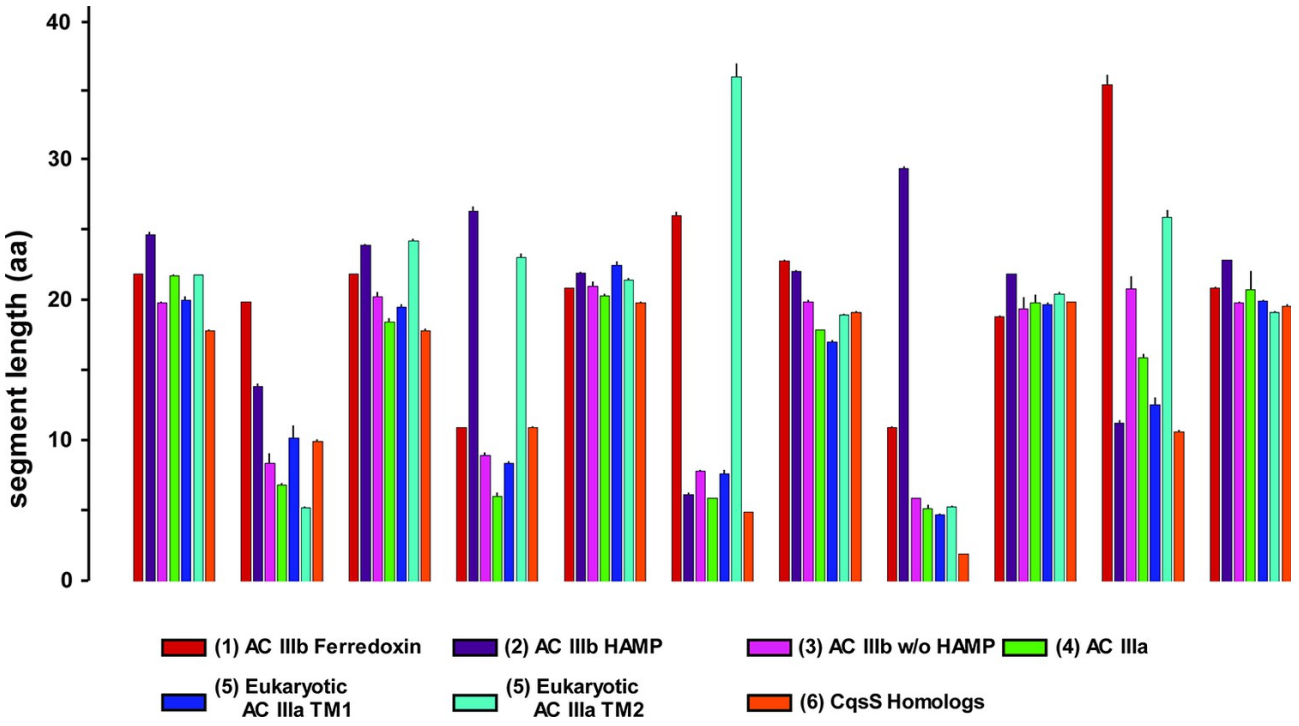


Fig. 7 supplement 1: Length comparisons of α-helices and loops of 6TM modules from adenylate cyclases and quorum sensors. Data are means ± S.E.M. of the sequence groups of all clusters shown in Figure 7A as indicated below. The color-coding used in Figure 7A has been conserved. The designations 'in' and 'out' indicate intra- and extracellular loops.

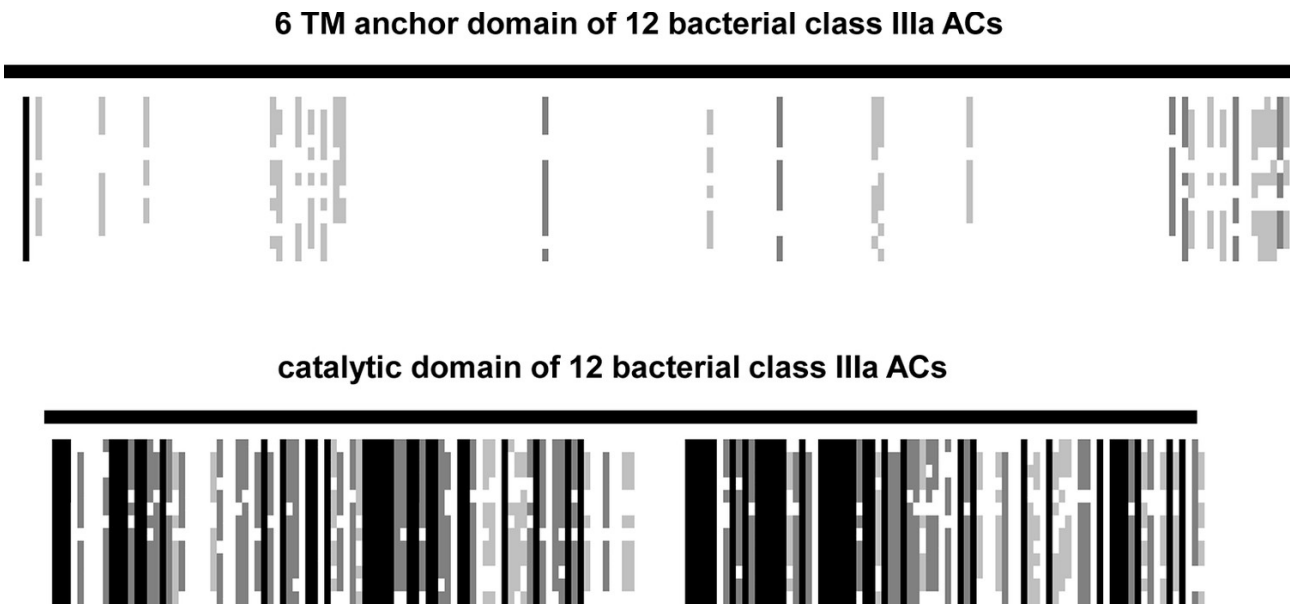


Fig. 7 supplement 2: The 12 sequences used for the alignment were from *Agrobacterium albertimagni,*WP_006725538; *Arthrospira maxima*, B5VUZ0; *Arthrospira platensis*, D5A5G2; *Beggiatoa*, A7BXS6; *Dechloromonas*, Q47AI8; *Hyphomicrobium*, C6QBG1; *Lyngbya*, A0YQ82; *Mesorhizobium*, LSHC420B00; *Microcoleus sp.*, PCC_7113; *Oscillatoria acuminata*, PCC_6304; *Nostoc*, YP_001866931; *Mycobacterium tuberculosis*, ALB18789 (Rv1625c). The alignment was made by ClustalW and adjusted and converted into a ‚bargraph ‘style alignment in Genedoc. Black lines reflect sequence identity, shades of grey different degrees of similarity, and white patches denote sequence diversity. Note the striking dissimilarity of the TM domains and the conservative nature of the catalytic domains.

B Manuscript Ziegler et al. (2017) FEBS J

Ziegler M., *Bassler J.*, Beltz S., Schultz A., Lupas A. N., and Schultz J. E.
Characterization of a novel signal transducer element intrinsic to class IIIa/b adenylate cyclases and guanylate cyclases
in: The FEBS Journal 284 (2017) pp. 1204-1217

Personal contribution:

Conception of the bioinformatic analysis (with ANL)

Acquisition of data shown in figures 3 and 5

Analysis and interpretation of this data (with ANL and JES)

Identification of the CTE in subclass IIIb and GCs

Conception and modeling of the mechanism shown in figure 8 (with ANL)

Drafting and revising the manuscript (with all authors)

The FEBS Journal

FEBS PRESS science publishing by scientists

Characterization of a novel signal transducer element intrinsic to class IIIa/b adenylate cyclases and guanylate cyclases

Miriam Ziegler[1], Jens Bassler[2], Stephanie Beltz[1], Anita Schultz[1], Andrei N. Lupas[2] and Joachim E. Schultz[1]

1 Pharmazeutisches Institut der Universität Tübingen, Germany
2 Max-Planck-Institut für Entwicklungsbiologie, Tübingen, Germany

Keywords
cyclic AMP; *Legionella*; quorum-sensing; receptor regulation; signal transduction

Correspondence
J. E. Schultz, Pharmazeutisches Institut der Universität Tübingen, Auf der Morgenstelle 8, D-72076 Tübingen, Germany
Fax: +49 7071 292476
Tel: +49 7071 2972475
E-mail: joachim.schultz@uni-tuebingen.de

(Received 27 October 2016, revised 9 January 2017, accepted 17 February 2017)

doi:10.1111/febs.14047

Adenylate cyclases (ACs) are signaling proteins that produce the second messenger cAMP. Class III ACs comprise four groups (class IIIa–d) of which class IIIa and IIIb ACs have been identified in bacteria and eukaryotes. Many class IIIa ACs are anchored to membranes via hexahelical domains. In eukaryotic ACs, membrane anchors are well conserved, suggesting that this region possesses important functional characteristics that are as yet unknown. To address this question, we replaced the hexahelical membrane anchor of the mycobacterial AC Rv1625c with the hexahelical quorum-sensing receptor from *Legionella*, LqsS. Using this chimera, we identified a novel 19-amino-acid cyclase transducer element (CTE) located N-terminally to the catalytic domain that links receptor stimulation to effector activation. Coupling of the receptor to the AC was possible at several positions distal to the membrane exit, resulting in stimulatory or inhibitory responses to the ligand *Legionella* autoinducer-1. In contrast, on the AC effector side functional coupling was only successful when starting with the CTE. Bioinformatics approaches established that distinct CTEs are widely present in class IIIa and IIIb ACs and in vertebrate guanylate cyclases. The data suggest that membrane-delimited receiver domains transduce regulatory signals to the downstream catalytic domains in an engineered AC model system. This may suggest a previously unknown mechanism for cellular cAMP regulation.

Introduction

Class III adenylate cyclases (ACs) are signaling proteins that produce the second messenger cAMP [1]. They are subcategorized into four groups, denoted class IIIa–d. Classes IIIa and IIIb are found in bacteria and eukaryotes, whereas classes IIIc and IIId are present in bacteria only [2]. Vertebrate genomes encode 10 AC isoforms, nine of which belong to class IIIa and are characterized by two different hexahelical transmembrane domains (6TM) and two complementary catalytic domains [3]. The pseudo-heterodimeric vertebrate ACs have monomeric bacterial homologs that require homodimerization for activity, such as the mycobacterial class IIIa AC Rv1625c, a prototype of the mammalian congeners [4]. The membrane anchors of such ACs comprise approximately 40% of the total protein. One feature of note is that they lack significant extramembranous loops connecting individual transmembrane spans [5]. Recently,

Abbreviations
AC, adenylate cyclase; CAI-1, cholera autoinducer-1; CTE, cyclase transducer element; DHp, dimerization and histidine phosphotransfer domain; GC, guanylate cyclase; LAI-1, *Legionella* autoinducer-1; QS, quorum-sensing.

membrane receptors with known ligands have been described that possess 6TM bundles with minimal-length α-helices and short connecting loops [6–8]. They are essentially identical in layout to the membrane anchors of 6TM ACs [3,4,9,10]. This receptor type is represented by bacterial quorum-sensing (QS) receptors present in *Vibrio*, *Legionella*, *Burkholderia* and other eubacteria [6,8,11,12]. The ligands for the QS receptors from *Vibrio*, CqsS, and *Legionella*, LqsS, have been identified as aliphatic hydroxy-keto lipids [8,11,12]. In the absence of sizeable extramembranous loops it is most likely that this type of ligand binds into the membrane space of the receptor protein.

Recently, we replaced the hexahelical membrane anchor of the Rv1625c AC by the QS receptor from *Vibrio* and demonstrated that the receptor conferred direct regulation by the extracellular ligand cholera autoinducer-1 (CAI-1) [5]. Short of the identification of a ligand for a membrane anchor of class IIIa ACs, this is a proof of principle for a receptor function of the 6TM AC anchors [5]. This assertion requires that the mechanism of intramolecular signal transduction from receptor to cytosolic effector in class IIIa ACs shares similarities with that in LqsS and CqsS histidine kinases [5,13].

So far, in membrane-delimited class IIIa ACs, cytosolic sequence elements that are potentially involved in signal transduction are unknown. Such elements should be located within the 60–80 residues that bridge the distance between the membrane exit and the start of the catalytic domain. Generating a chimera between the QS receptor from *Legionella pneumophila* and the Rv1625c AC, we asked: What are the sequence requirements for a functional coupling? Here we identify and characterize a conserved segment of 19 amino acids, the cyclase transducer element (CTE), which is indispensable for signal transduction and which had escaped detection in the past. Bioinformatic analyses demonstrate that CTEs are present in class IIIa and IIIb ACs, and also in ligand-regulated guanylate cyclases (GCs) from vertebrates.

Results

Connecting the QS receptor LqsS to Arg218 of the Rv1625c AC

Recently, we have demonstrated that CqsS, the QS receptor from *Vibrio*, can replace the 6TM membrane anchor of the Rv1625c AC [5]. AC activity in such a CqsS–Rv1625c chimera is directly stimulated 190% by 1 μM of the ligand CAI-1 ((*S*)-3-hydroxy-tridecan-4-one; [5]). Here, we replaced the 6TM bundle of Rv1625c AC with the QS receptor from *Legionella*, LqsS, and initially used the equivalent sequence positions in LqsS (Met191) and in Rv1625c (Arg218) for linking (Fig. 1). AC activity was stimulated in the chimera by the *Legionella* autoinducer-1 (LAI-1; (*S*)-3-hydroxy-pentadecan-4-one). However, its maximal stimulation did not exceed 120%. Half-maximal activation (EC_{50}) was graphically estimated to be 48 nM LAI-1 (Fig. 1). The rather weak, albeit statistically significant stimulation might have indicated that Rv1625c partners with LqsS less efficiently than with the CqsS receptor from *Vibrio*, or that the point of connection between LqsS and Rv1625c was suboptimal for intramolecular signal transduction.

Therefore, LqsS was joined with Rv1625c at 11 different positions distal to the predicted cytosolic exit of transmembrane helix 6 at Tyr178 while the point of connection to the cyclase, Arg218, remained unchanged. The results were unexpected (Fig. 2A). Basal activities ranged from 15 to 64 nmol cAMP·mg^{-1}·min^{-1}, comparable to the activity of the unabridged membrane-

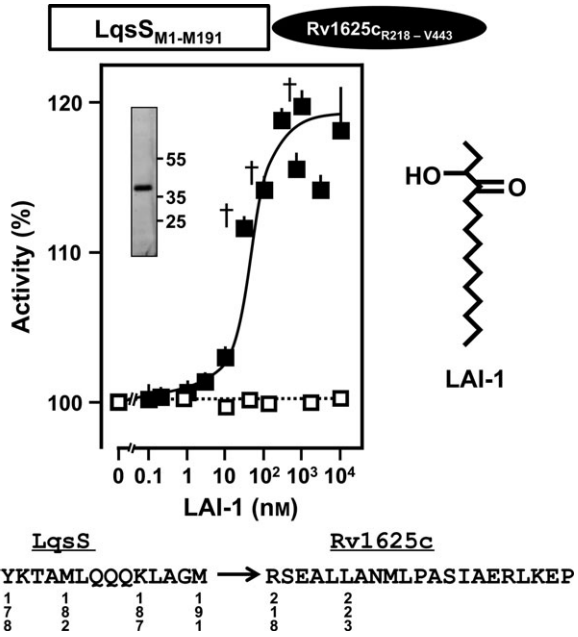


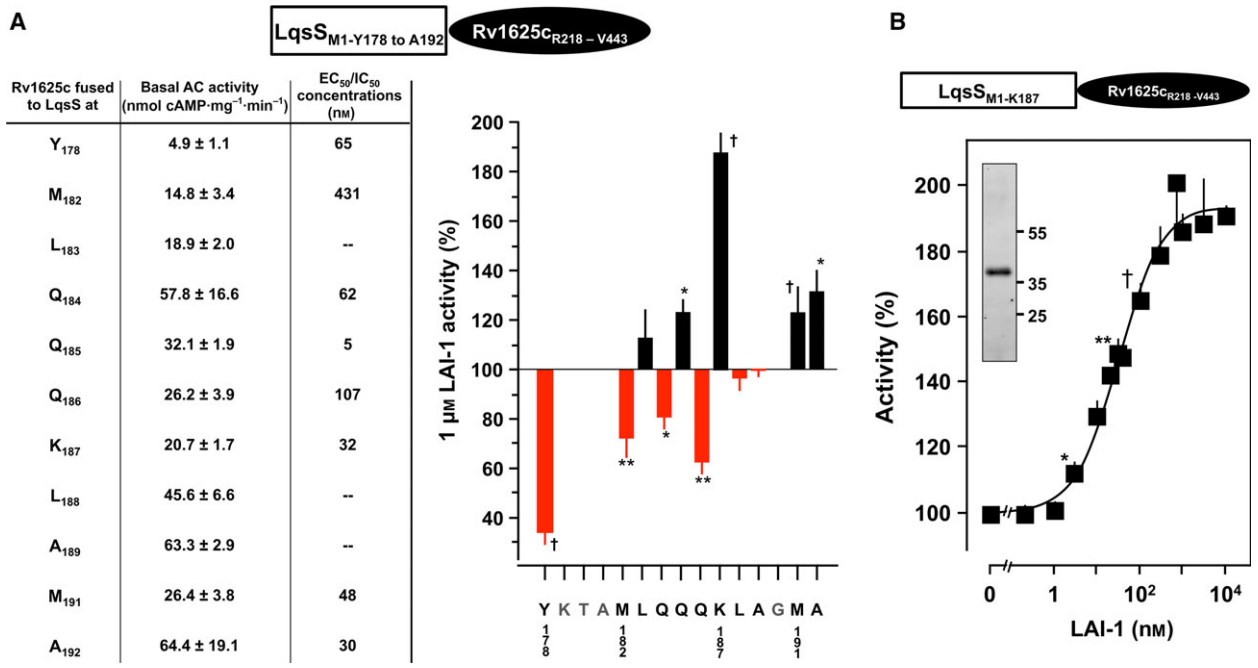
Fig. 1. Stimulation of the chimera LqsS–Rv1625c by the QS ligand LAI-1. The scheme for the chimeric construct is shown at the top, sequence information of the critical junction point at the bottom, with respective numbering for LqsS and Rv1625c. The basal activity (100%) was 26.4 nmol cAMP·mg^{-1}·min^{-1}. The EC_{50} concentration was estimated to be 48 nM. $^{\dagger}P < 0.001$. For clarity further significances are omitted. The catalytic domain of Rv1625c$_{R218-V443}$ (14.3 nmol cAMP·mg^{-1}·min^{-1}) was not affected by 10 μM LAI-1 (open squares). The inset shows the western blot of the protein. Results are presented as means ± SEM ($n \geq 5$). At the right is the structure of the ligand.

anchored Rv1625c AC [4]. Levels of protein expression were equivalent for these constructs, as analyzed by western blotting (data not shown). The chimera linked at Tyr178, i.e. directly at the presumed C terminus of TM6, was only modestly active (4.9 nmol cAMP·mg^{-1}·min^{-1}). However, with almost 70% inhibition, it was significantly affected by LAI-1 (estimated IC_{50} = 65 nM; Fig. 2A). Inhibition of chimeras was also observed in three other constructs conjoined four, six and eight amino acids distal to the membrane exit of LqsS, respectively (at Met182, Gln184 and Gln186; Fig. 2A). Notably, in the latter chimeras, basal AC activities were considerably higher. This may indicate that formation of a productive AC dimer was facilitated with greater distance from the predicted membrane exit. Significant stimulation was observed at connection positions more distal to the membrane exit, such as Gln185, Lys187, Met191 and Ala192 (Fig. 2A). Maximal stimulation of 190% was achieved in the $LqsS_{M1\text{-}K187}$–$Rv1625c_{R218\text{-}V443}$ chimera (Fig. 2B). The estimated EC_{50} was 48 nM LAI-1.

Taken together, mostly inhibitory responses were observed when the point of connection was within eight residues of the membrane exit, whereas stimulatory responses prevailed with longer linker sequences (Fig. 2). The estimated ligand concentrations required for half-maximal responses were in the same order of magnitude for all constructs (respective concentration–response curves not shown). This indicated that the modifications did not significantly affect receptor–ligand interactions. Lys187 in LqsS was used as the point of connection in subsequent constructs.

Connecting the AC Rv1625c to Lys187 of the QS receptor LqsS

Next, we analyzed the connecting point on the side of the Rv1625c AC. Inspection of the linker sequences between membrane domains and catalytic domains of type IIIa ACs indicated a conserved segment characterized by an invariant central proline. Its likely C terminus was evident from a distinct drop of sequence conservation. The N terminus could not be identified as clearly (Fig. 3). Earlier experiments had successfully used Arg218 of Rv1625c as the fusion point for chimeras between the AC and CqsS [5]. Similarly, the



Rv1625c fused to LqsS at	Basal AC activity (nmol cAMP·mg^{-1}·min^{-1})	EC_{50}/IC_{50} concentrations (nM)
Y_{178}	4.9 ± 1.1	65
M_{182}	14.8 ± 3.4	431
L_{183}	18.9 ± 2.0	--
Q_{184}	57.8 ± 16.6	62
Q_{185}	32.1 ± 1.9	5
Q_{186}	26.2 ± 3.9	107
K_{187}	20.7 ± 1.7	32
L_{188}	45.6 ± 6.6	--
A_{189}	63.3 ± 2.9	--
M_{191}	26.4 ± 3.8	48
A_{192}	64.4 ± 19.1	30

Fig. 2. The sign of the output signal depends on the point of connection of the LqsS-receptor to the Rv1625c AC. (A) Left, basal AC activities of chimeras with different connection points of LqsS and respective EC_{50}/IC_{50} values graphically derived from corresponding concentration–response curves that were carried out for each particular point of connection using at least five different LAI-1 concentrations from 1 to 1000 nM. Right, response of the 11 chimeras to 1 µM LAI-1. Residues 179–181 and 190 (grey typeset) were not tested. (B) LAI-1 concentration–response curve for the LqsS construct connected to Rv1625c via Lys187. Data are normalized to the basal activity given in (A). The inset show a western blot (1 µg per lane) of the expressed protein indicating absence of proteolysis. Results are presented as means ± SEM ($n \geq 3$). Significances: *$P < 0.05$; **$P < 0.01$; †$P < 0.001$ compared with respective basal activities. To avoid cluttering the figure, not all significant points are marked.

corresponding position in the class IIIa AC CyaG from *Arthrospira platensis* (Arg456) was successfully fused to the 2TM chemotaxis receptor Tsr from *Escherichia coli*, resulting in a serine-sensitive AC chimera [14]. We tested LqsS-Lys_{187}–Rv1625c constructs connected at up to four additional amino acids prior to Arg218 (residues 214–217 in Fig. 4). The resulting chimeras had high basal activities, but were unresponsive to LAI-1 (Fig. 4). Connection via Arg218 resulted in a 75% drop of basal activity, but highly significant stimulation by LAI-1 (Figs 2 and 4). The importance of Arg218 in Rv1625c seemed contrary to the high sequence variability at this position in homologs (Fig. 3). To clarify its functional role, we generated three point mutants, replacing the positively charged Arg218 with a hydrophobic (R218L), a polar (R218Q), or an acidic (R218D) residue. In all instances, the stimulation of the chimeric protein and ligand potency were comparable to the Arg218 parent construct (190% stimulation), i.e. 210% for R218L, 241% for R218Q and 206% stimulation for R218D. This indicated that stimulation was dependent on the length of this segment rather than on the physical properties of this particular amino acid.

Stepwise deletion of Arg218 and six downstream residues (Ser219 to Ala224) resulted in a progressive loss of AC activity and regulation (Fig. 4). The outright deletion of 19 residues downstream of Arg218 in Rv1625c, up to Lys236, resulted in low basal AC activity (Lys236 in Fig. 4; 170 pmol $cAMP \cdot mg^{-1} \cdot min^{-1}$) and a loss of regulation. This indicated that residues Arg218 to Lys236 have an essential function in the regulation of Rv1625c AC activity. We termed this region the CTE for its putative role in intramolecular signal transduction.

Bioinformatic characterization of a universal cyclase transducer element

Asking whether the CTE is in fact a novel signaling element, rather than a constitutive component of the cyclase fold, we comprehensively analyzed its presence in sequences of nucleotide cyclases by bioinformatics. We found it N-terminal to catalytic domains of class IIIa and IIIb ACs, including the membrane-bound and soluble isoforms of vertebrates (Fig. 3). CTEs were absent in class IIIc and IIId ACs. CTEs were also present in the ligand-regulated eukaryotic GCs, which are closely related to class IIIa ACs, but not in diguanylate cyclases, which form an outgroup to class III ACs. Crystal structures of CTEs exist from a GC from rat, where it was solved together with the upstream signaling helix [S-helix; Protein Data Bank (pdb) : 3hls], and of the class IIIb soluble AC from human (hAC10), where both CTEs of the pseudo-heterodimer were solved together with the catalytic domains (pdb: 4clf; [15,16]). We conclude that the CTE is a distinct element, which likely evolved in the common ancestor of modern class IIIa and IIIb cyclases.

We noted that the CTE sequences are to some extent divergent (Fig. 3). The invariant proline, which had originally caught our attention, is conserved only in CTEs of class IIIa ACs. In the class IIIb AC10 from vertebrates, this proline is conserved in CTEs of the first catalytic (C1) domain, whereas it is absent in CTEs of the C2 domain. It is also absent in CTEs of the homodimeric bacterial class IIIb ACs. Secondary structure predictions, however, are highly similar for all CTEs, irrespective of the presence of this proline residue. This is in agreement with the aforementioned crystal structures, which superimpose to less than 1.5 Å root mean square deviation, even though the C-terminal CTE of the hAC10 is divergent and lacks the central proline. In pseudo-heterodimeric ACs from vertebrates, CTEs of C1 catalytic domains contain an invariant serine three residues upstream of the central proline, whereas this position is occupied by an invariant asparagine in C2 CTEs (Fig. 3). CTE sequences from bacterial ACs and eukaryotic GCs are intermediate, having either serine or asparagine, and probably reflect the ancestral state.

Prompted by this speciation we conducted a cluster analysis of CTEs from class IIIa ACs. The sequences separated, as expected, into three groups along their association with bacterial ACs, vertebrate C1 or vertebrate C2 domain (Fig. 5). The vertebrate groups were further subdivided according to their origin from the nine membrane-delimited isoforms, due to minor, yet conserved differences between their sequences, closely mirroring a cluster map of the catalytic domains (data not shown). This implies that CTEs must have coevolved together with their respective catalytic domains and membrane anchors because the isoform designations were originally based on sequence variations in the transmembrane domains [1,2].

Exchange of transducer elements between adenylate cyclase isoforms

The close coevolution of CTEs with their cognate catalytic domains prompted us to ask whether, in the homodimeric LqsS–Rv1625c chimera, the CTE might be functionally replaceable by either of the two CTEs of the mammalian AC5 isoform (Fig. 6). In all cases, introducing an AC5 CTE into the LqsS–Rv1625c chimera resulted in a dramatic drop in basal AC activity

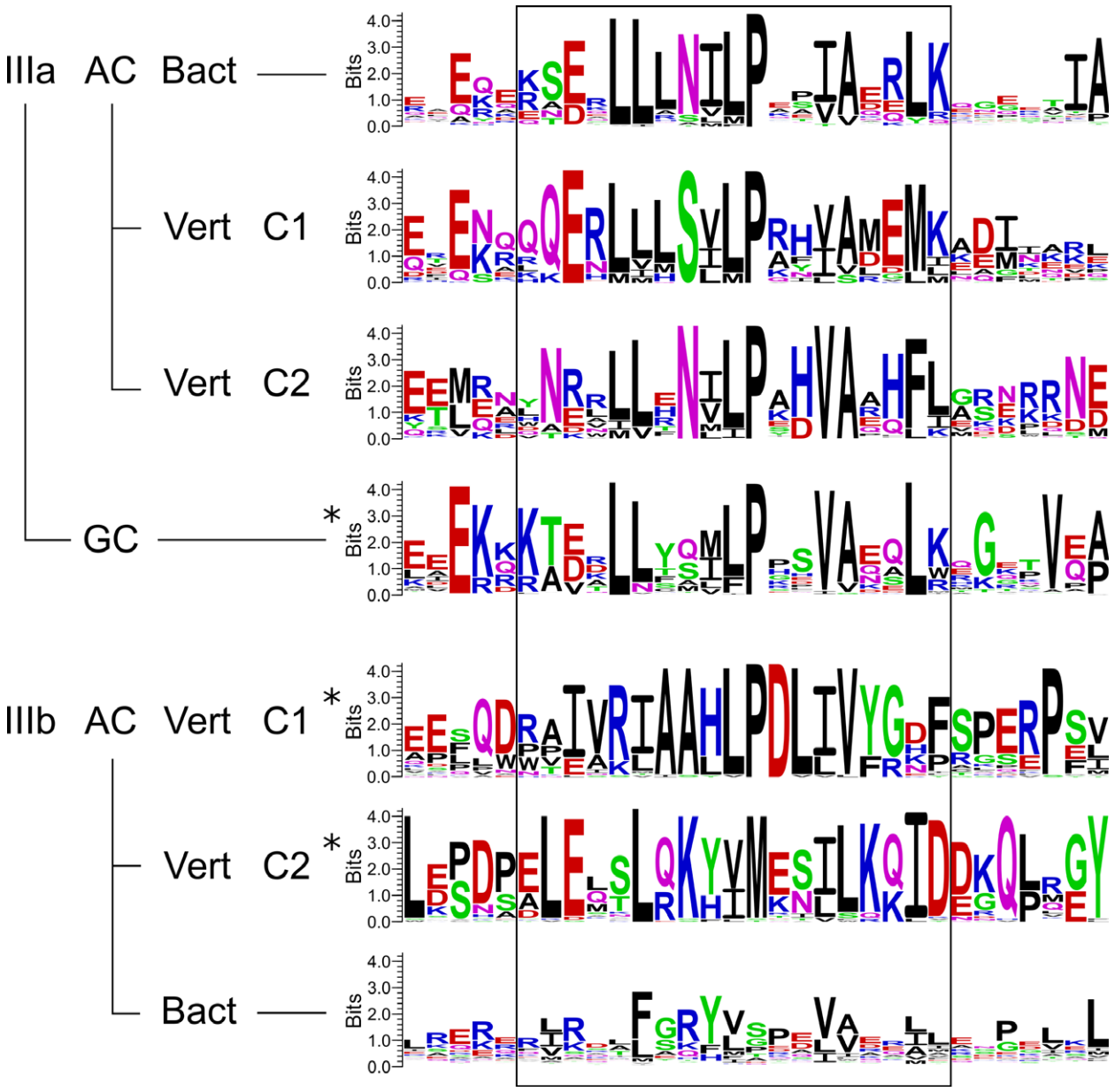


Fig. 3. Sequence diversity of the CTE. The sequence logos are grouped according to classification, phylogeny and origin of the catalytic subunit. Bacterial ACs are labeled with 'Bact', vertebrate ACs with 'Vert'. For vertebrate pseudoheterodimers, C1 and C2 denote the catalytic domains to which the CTE is coupled. Vertebrate class IIIb ACs represent exclusively the soluble AC10 isoform. Note that all subgroups show specific sequence features, most prominently the presence or absence of a central invariant proline. Asterisks denote groups with a solved crystal structure. The following numbers of sequences were used in preparation of the logos: IIIa bacterial: 314; IIIa vertebrate C1: 469; IIIa vertebrate C2: 482; IIIa GCs: 624; IIIb vertebrate C1: 38; IIIb vertebrate C2: 51; IIIb bacterial: 1008. The data set is available at Figshare.

(Fig. 6A). This emphasized the close relationship between catalytic domains and their cognate CTEs. Nevertheless, we noted that the CTE from AC5-C1 passed the LAI-1 signal to the catalytic domains, both as a homodimer and as an engineered heterodimer in conjunction with the Rv1625c CTE, resulting in significant stimulation. All constructs in which the CTE from AC5-C2 was involved were barely or not at all responsive to LAI-1 (Fig. 6A). We propose that CTEs are critically involved in controlling the formation of an active AC complex between the catalytically inactive monomers. To explore whether this depended

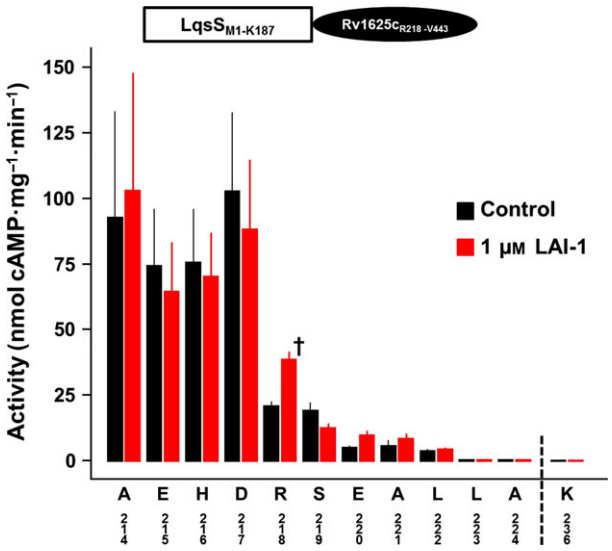


Fig. 4. The basal activities (black bars) and the activities at 1 μM LAI-1 (red bars) are shown for the denoted points of connection of Rv1625c to Lys187 from LqsS. Only the parent construct LqsS$_{M1-K187}$–Rv1625c$_{R218-V443}$ was significantly stimulated ($^{\dagger}P < 0.001$) by LAI-1. See Table S1 for sequence information. Results are means ± SEM ($n \geq 2$).

on the presence of the membrane domains, we employed a soluble Rv1625c construct, linked by an established tetradecapeptide used previously for this purpose [4,17]. In these soluble constructs, the catalytic domains are mutated such that hetero-association is required for activity [4]. The affinity-purified Rv1625c $AC_{204–443}$ heterodimer had high AC activity (380 nmol cAMP·mg^{-1}·min^{-1}; Fig. 6B). Replacement of the Rv1625c CTE by the human AC5 CTEs reduced AC activity by 80–95% compared with the parent construct (Fig. 6B). This implies that even without membrane attachment, formation of an active catalyst apparently required a delicate CTE interaction. In the context of the Rv1625c catalytic domains, the conditions for a seamless interaction of the two AC5 CTEs are missing, and the formation of the Rv1625c catalytic homodimer was impaired. This finding may be relevant concerning the enormous difficulties in the past to obtain soluble active AC dimers from a single AC isoform. Mostly, dimers composed of C1 and C2 domains from different vertebrate AC isoforms were used for enzymatic assays and for crystallization [17–22]. As a role for the CTEs was not appreciated at the time they were generally not included in these constructs.

In this context it is further notable that functional coupling of LqsS to the Rv1625c AC via the Rv1625c CTE was partly possible although the His-kinase from which the LqsS receptor domain was derived has no CTEs. To us, this suggested that CTEs operate as cyclase-specific signal transducers, which convert incoming signals into conformational motions that regulate the dimerization of the catalytic domains (see below).

Interaction of the cyclase transducer element with the S-helix

With the CTE as an independent transducer element it should be possible to connect it functionally to other independent signal transducing domains. The S-helix is a conserved coiled-coil domain of up to 50 residues, found between receiver and output domains in diverse sensory proteins, including histidine kinases and nucleotide cyclases [23,24]. In the retinal GC an S-helix is located directly in front of the CTE [15]. In a construct connecting the receiver domain of the *E. coli* serine receptor Tsr via a HAMP domain, a CTE and an S-helix to the catalytic domain of the AC CyaG from *A. platensis*, robust regulation by serine was observed, i.e. all participants interacted productively [14,25]. Removal of the S-helix inverted the sign of the output signal [14].

To investigate whether the CTE of Rv1625c would functionally interact with a heterologous S-helix, we incorporated the 25-amino-acid-long S-helix of CyaG into the LqsS–Rv1625c chimera (Fig. 7). These 'triple' chimeras, consisting of the LqsS receptor, the S-helix from CyaG and the CTE and cyclase domain of Rv1625c, were well expressed in *E. coli* (see western blots in Fig. 7). In the protein with the S-helix in front of the CTE, the sign of the receptor stimulation was inverted compared with the parent construct, i.e. LAI-1 inhibited AC activity by almost 50% in a concentration-dependent manner (Fig. 7). The estimated ligand concentration required for a half-maximal inhibition (48 nM) was similar to the EC_{50} for half-maximal stimulation of the parent construct (21 nM; Fig. 7). This means that the ligand–receptor interactions remained unimpaired by insertion of an additional transducer element.

As a control, the length of the S-helix was mimicked by a respective extension of LqsS by 25 residues (LAGMAAAAGMIAHELRSPLLGIKSG) in front of the CTE. This stretch is predicted as helical but shares no sequence similarity with the S-helix. This led to an unregulated chimera, which had a basal AC activity of 0.58 nmol cAMP·mg^{-1}·min^{-1} (Fig. 7). As a second control, the S-helix was inserted posterior to the CTE. AC activity was considerably diminished in the chimeric protein, although expression of the chimera was equivalent to the other constructs (see western blots in Fig. 7). On the basis of the attenuated basal activity LAI-1

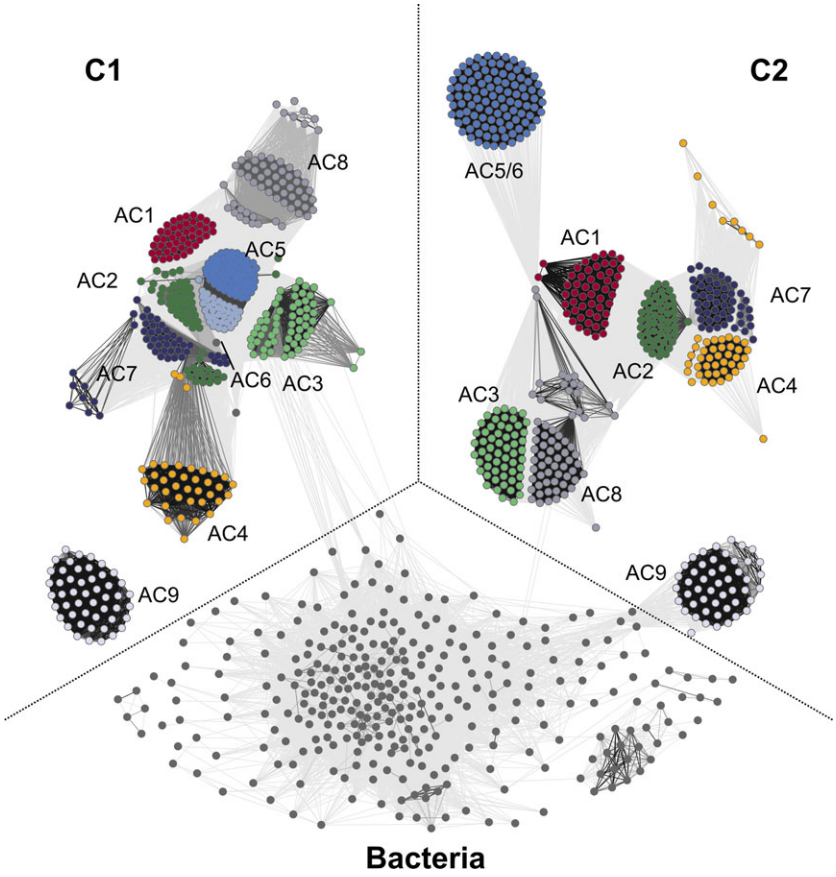


Fig. 5. Cluster map of CTEs. The sequences of bacterial and vertebrate class IIIa CTEs (Fig. 3) were analyzed using CLANS (*P*-value cutoff 1×10^{-10}, attraction value 0.05 and repulse value 0.5). Each dot represents a single sequence. Above-threshold hits are shown as connecting lines. Darker line color indicates better pairwise BLAST hits. The three sectors are partitioned by broken lines. Cluster labeling indicates the AC isoform. The segregation shows that CTEs from vertebrate class IIIa ACs are highly specific for the C1- and C2-domain origins as well as the peculiar AC isoforms. Bacterial sequences display no distinct clustering pattern. The data set comprising a total of 1265 AC sequences that were used for the cluster analysis is available at Figshare.

stimulated AC activity with an identical potency as the parent construct (Fig. 7). Taken together, the data demonstrated that the effect of ligands on the output domain was dependent on the order of the transducer elements. Similarly, the data support the notion that the CTE represents a signal transducer element capable of functional interaction with other transducer elements. The interaction appears to be independent of the nature of the membrane domain [14,25].

A structural model for CTE signal transduction

Next, we wondered how CTEs might mechanistically transduce the signal. Our considerations were guided by the conspicuous parallels between the CTE and the analogous transducer element of histidine kinases, the dimerization and histidine phosphotransfer domain (DHp). Both are frequently connected at their N-termini to signal transduction domains, such as HAMP, GAF, PAS, or S-helices, from which they receive their input [23,26,27]. Both elements invariably connect directly at their C-termini to the catalytic output domains, whose level of activity they set. In DHp, setting of the output, i.e. phosphatase *versus* kinase activity, has been shown to involve transitions between symmetric and asymmetric conformations [26,28]. We envisage similarly concerted transitions between nearly isoenergetic states for the CTE. Given the diversity of domains capable of conveying signals to both DHp and CTE, it appears that these efficiently convert different conformational inputs into defined structural changes that determine the activity of the output domains. Polar linkers appear to be a key feature in this process connecting the upstream signaling domains to the transducer elements [26,28]. In DHp, mutations in this linker region are known to severely impair signaling [14]. Accordingly, we note the widespread presence of such linkers N-terminal to the CTEs of type IIIa and IIIb ACs and GCs. The functional analogies in signaling between CTEs and DHp provide a rationale for our ability to generate chimeras between class IIIa ACs and His kinases.

Based on this analogy, we have considered possible conformational changes of CTEs during signal transduction. The available structures of CTEs show two helices, separated by a central bend of approximately 45° [15,16]. In the pseudo-heterodimeric CTE of human AC10 the N-terminal helices form a short

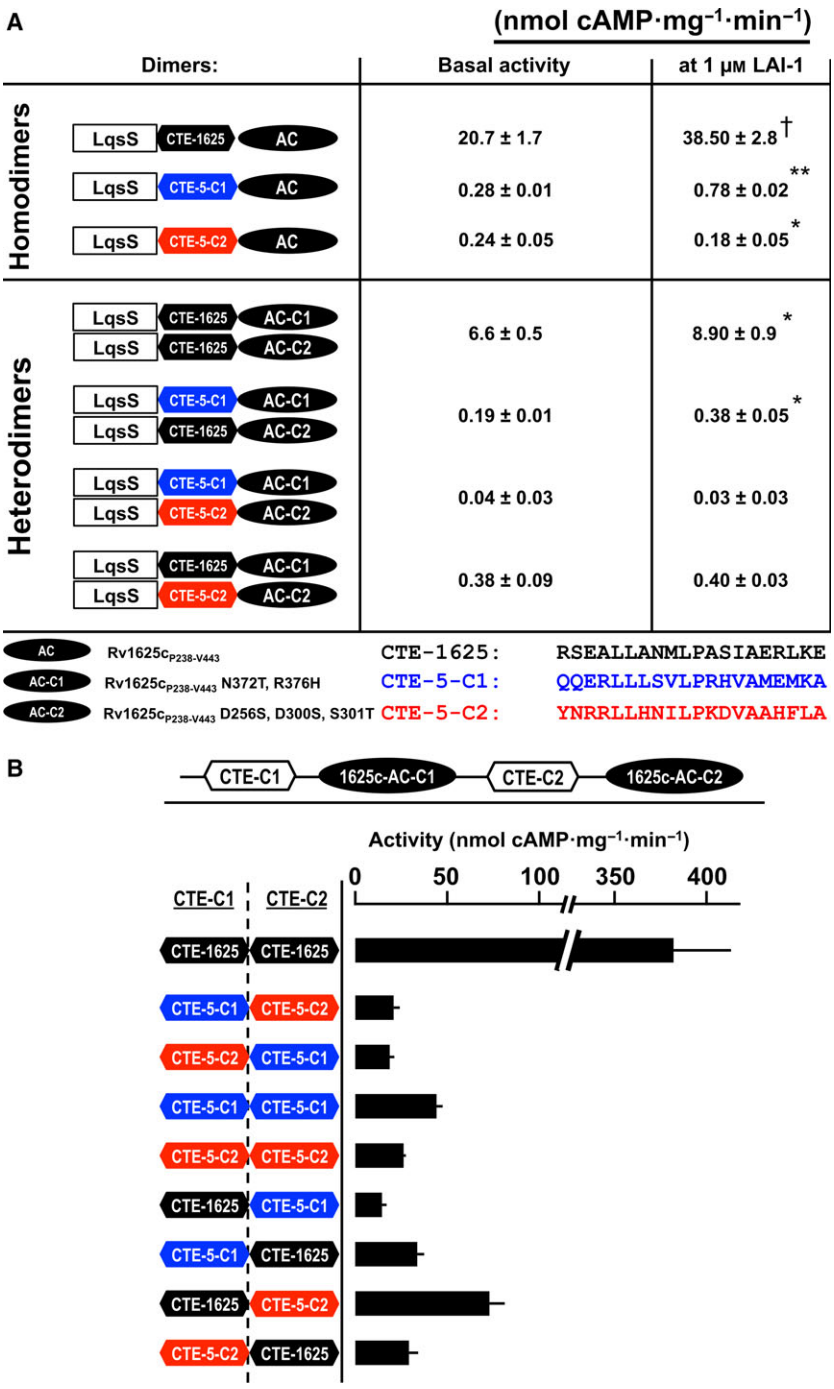


Fig. 6. Replacement of the Rv1625c CTE with those from human AC5-C1 and/or -C2. (A) Homodimers: note the drop of activity upon insertion of hAC5 CTEs in LqsS–Rv1625c AC chimeras. Heterodimers: by introduction of inactivating, yet complementing, mutations into Rv1625c (AC-C1 and AC-C2; [4]) exclusively the desired heterodimers can form a catalytic center. White box, LqsS receptor domain M1-K187; black icons, domains from Rv1625c AC; blue icons, CTE from AC5-C1; red icons, CTE from AC5-C2. The respective CTE sequences are shown at the lower right. Results are means ± SEM ($n \geq 3$; *$P < 0.05$; **$P < 0.01$; †$P < 0.001$). (B) Soluble heterodimers of AC Rv1625c with CTE replacements from human AC5-C1 and AC5-C2 as controls. Each replacement by a CTE from AC5 resulted in a statistically significant ($P < 0.001$) drop of activity of 80–95%. Results are means ± SEM ($n \geq 6$). Sequence information is available in Table S1.

coiled coil, from which the C-terminal helices extend into opposite directions. Thus, the CTE is directly compatible with upstream domains that end in a coiled-coil domain, notably HAMP and the S-helix. The hydrophobic register of the N-terminal helices extends with only minor disturbances into the C-terminal helices, suggesting that the CTE could switch between bent and straight conformations (Fig. 8). In the future this might be amenable to experimental testing by introduction of appropriate mutations.

An analogous structural transition was reported for the pH-sensing Rv1264 AC from *Mycobacterium tuberculosis* [29], which, being a class IIIc AC, lacks a CTE. In the active state, the segment N-terminal to the catalytic domain constitutes a partly unwound α-helix and the catalytic domains are free to self-

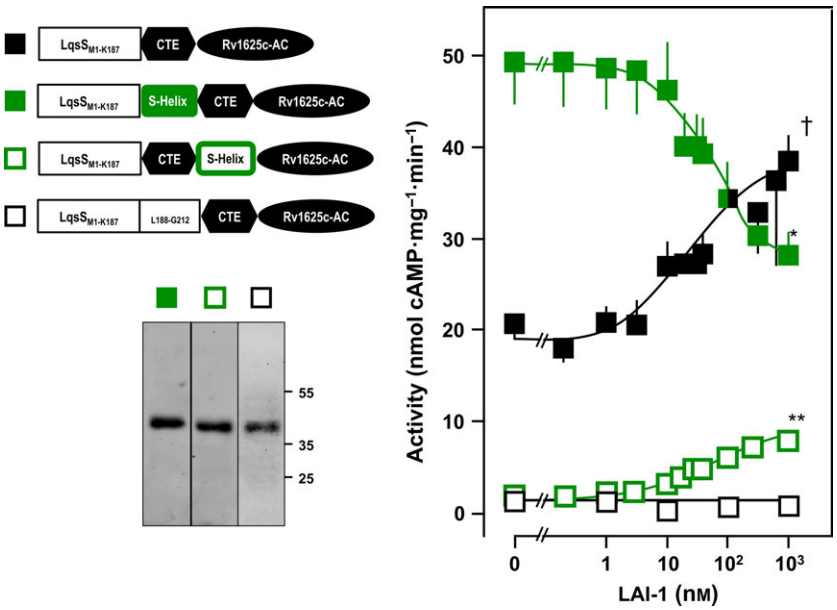


Fig. 7. The S-helix interacts with the CTE and inverts the sign of the signal. Filled green squares, insertion of the S-helix N-terminal of CTE; filled black squares, parent chimera replotted from Fig. 2B for a direct comparison (estimated EC_{50} = 21 nM; for western blot see Fig. 2B). Controls: open green squares, S-helix C-terminal of CTE (estimated EC_{50} = 34 nM); open black squares, extension of LqsS by 25 residues to Gly212. Results are means ± SEM ($n \geq 2$). *$P < 0.05$; **$P < 0.01$; †$P < 0.001$. The design of the constructs is indicated at the upper left. Western blots of respective constructs at lower left. See Table S1 for sequence information.

assemble into an active dimer. In the inactive state the N-terminal segment forms a rigid helix that holds the inactive monomers apart [29]. Similarly, the CTE might affect the conformational flexibility needed by the catalytic domains to assemble the active dimer, suggesting that the straight conformation of CTE correlates with the inactive state. The pairwise adaptation required for such conformational switching provides an attractive explanation for the coevolution of membrane domains, CTEs and catalytic domains (Figs 3 and 5). In this context, the differences between the CTEs of pseudo-heterodimers are particularly striking, because they suggest that, in vertebrates, CTEs may have diverged to enable the decoding of more complex, possibly asymmetric signals.

Discussion

The hexahelical membrane anchors of class IIIa ACs, in particular of the nine mammalian isoforms, have remained a sort of biochemical mystery for the past 27 years. At the time of the initial sequencing of a mammalian AC, a transporter or channel function was proposed [9]. At present, we know that the membrane anchors in eukaryotic ACs are conserved in an isoform-specific manner from coelacanth to man, separated by approximately 400 million years of evolution [5]. However, a defined role beyond that of a passive membrane anchor has remained elusive. Most recently, we have provided evidence for a role as a ligand receptor by biochemical characterization of a chimera between the QS receptor from *Vibrio harveyi*, CqsS, and the Rv1625c AC. A bioinformatic analysis of shared properties of the membrane domains supported the proposed receptor function [5]. For a receptor to transmit an extracellular signal to a cytosolic effector domain, a suitable molecular signal transducer or converter for a shared AC effector domain is likely to be required. Such a molecular transducer entity present in class IIIa and IIIb ACs has been identified and characterized in this study.

Examination of the connecting points between the LqsS receptor and the AC effector domain identified a conserved segment of about 19 residues present in class IIIa and IIIb ACs. It is termed CTE. Probably, a centrally located proline is important for structure and function. The functional interaction between CTE and the S-helix bolsters the claim that CTEs are signaling devices involved in regulation of class IIIa and b ACs. This function probably is not restricted to membrane-delimited ACs but may well operate in cytosolic ACs composed of receiver and effector domains [2,13]. The sequence differences between CTEs probably are functionally significant (Fig. 6A). This indicates that the receptor, the CTE and the effector output domain communicate in a concerted manner reminiscent of bacterial signaling through HAMP domains [30,31]. Most likely, the receptor–CTE–effector trio is tuned to each other such that the small energetic differences caused by ligand binding to the sensory domain are scaled to trigger the conformational changes required in the respective enzymatic output domain. Such mechanisms provide a high degree of signaling specificity.

On the other hand, one may ask whether the CTEs could not play a rather passive role in the allosteric connection with a sensor domain and that the catalytic

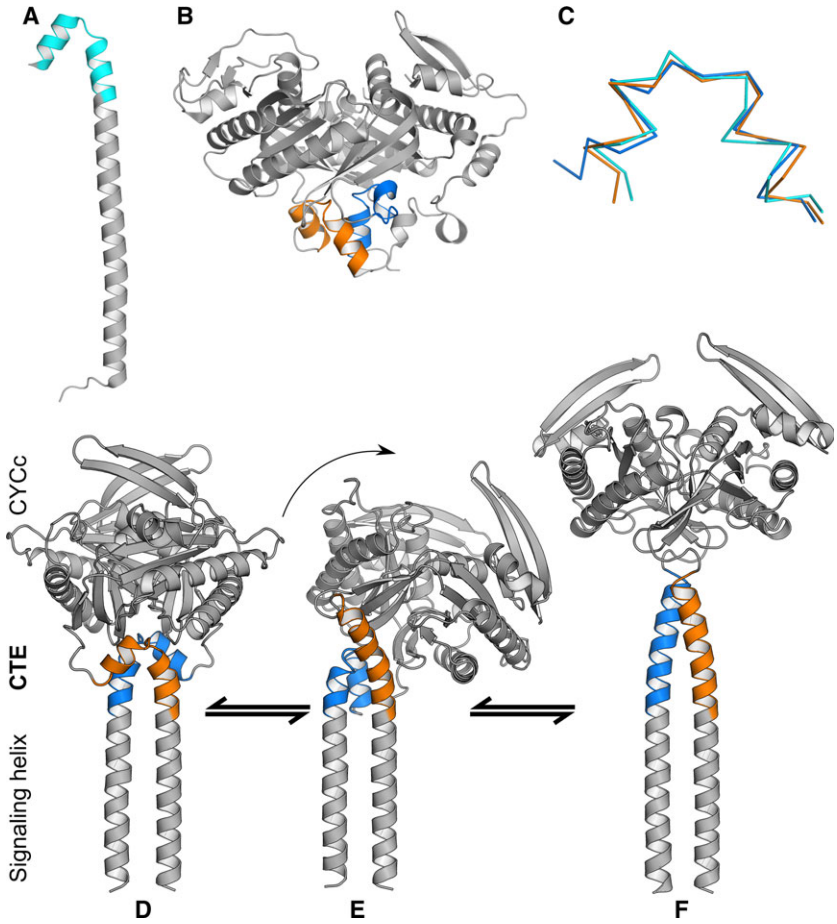


Fig. 8. Top: CTEs in available crystal structures. (A) Structure of a CTE from a rat guanylate cyclase together with the upstream signaling helix (S-helix) (pdb: 3hls). (B) Structure of the class IIIb soluble AC from human (AC10), where both CTEs of the pseudo-heterodimer were solved together with the catalytic domains (pdb: 4c lf) [16]. (C) Superposition of the CTEs in (A) and (B). The CTEs are structurally highly similar and superimpose at a root mean square deviation of less than 1.5 Å. Bottom: a structural model for signal transduction through CTEs. It is proposed that the CTE operates as a hinge that can adopt a bent or a straight conformation. (D) A computational model of the S-helix-CTE-cyclase constructs based on the structures in (A) and (B). The CTEs (highlighted in blue and orange) are in a bent conformation. (E) The transition progresses through an asymmetric state in which one CTE has adopted a hypothetical straight conformation while the other remains bent. (F) Both CTEs adopt the hypothetical straight conformation. The hydrophobic residues in the C-terminal helices interact, thereby extending the coiled-coil formed by the N-terminal helices.

domains and CTEs are 'finely tuned' to each other for the purpose of catalysis only, but not allosteric signal transmission between the ligand-binding receptor and the catalytic dimer. We consider this a rather remote possibility because sequence and structural information indicates essentially identical catalytic mechanisms in all class III ACs [18,29,32,33]. In addition, it would be difficult to rationalize the reasons for the high degree of isoform-specific evolutionary conservation, which is apparent in CTEs from vertebrate ACs (Fig. 5). Remarkably, in soluble Rv1625c proteins without CTE specific activities drop by up to 90% [4] whereas in soluble mammalian AC-5C1:AC-2C2 heterodimeric constructs, in which the CTEs are absent, basal AC activity is low, but synergistically stimulated by forskolin/Gsα (~ 10 μmol·min^{-1}·mg^{-1}; [34]). However, the failure to functionally combine soluble C1 and C2 catalytic domains from a single vertebrate isoform may actually suggest a conformational inability to properly interact in the absence of the respective CTEs. Finally, the functional interaction between the CTE and the established S-helix (Fig. 7) strongly implies an intrinsic signal transducing capacity of CTEs.

In summary, the data provide a stringent argument that membrane-delimited ACs will turn out to be primary signal receivers. For the nine membrane-delimited isoforms of vertebrates, this implies that next to the well-characterized indirect regulation by G-proteins, yet unknown direct regulatory signals impinge on the AC membrane domains. This will establish an entirely new regulatory pathway of cAMP biosynthesis. This argument is further supported by recent reports of two mutations in the human AC type 5 that cause familial dyskinesia and facial myokymia. The point mutations are located between the hexahelical membrane anchors and the respective CTEs [35–37]. These single point mutations do not affect AC activity in HEK293 cells transfected with wild-type or respective AC5 mutants, indicating that expression or protein folding is unimpaired. Along the model proposed here, it appears plausible that an as yet uncharacterized mode of AC regulation is dysfunctional in these mutants. A similarly located mutation is also known in the human retinal GC, causing autosomal dominant cone–rod dystrophy [38,39].

Experimental procedures

Materials

A plasmid containing the LqsS His-kinase receptor DNA (Uniprot entry Q5ZRY7) was a gift from H. Hilbi (LMU, Munich, Germany). *Mycobacterium tuberculosis* AC Rv1625c (Uniprot entry P9WQ35) and mutant clones with point mutations in catalytic residues were available in the laboratory [4]. Radiochemicals were from Hartmann Analytik (Braunschweig, Germany) and Perkin Elmer (Waltham, MA, USA). All enzymes were purchased from New England Biolabs (Ipswich, MA, USA) or Roche Diagnostics (Basel, Switzerland). Other chemicals were from Sigma-Aldrich (St. Louis, MO, USA), Roche Diagnostics, Merck (Darmstadt, Germany) and Roth (Karlsruhe, Germany). LAI-1 was synthesized in-house according to [12] and dissolved in DMSO. Because the ligand has surfactant properties, concentration–response curves usually were run up to 1 μM LAI-1. The 25-amino-acid S-helix was from the *A. platensis* AC CyaG (Uniprot entry K1VRL5; A_{431}LENTNRELEQRVLERTAALLQEKE$_{455}$). For the CTE replacements in the Rv1625c AC (Arg218 to Glu237) the corresponding 20-amino-acid segments from human AC5-C1 and AC5-C2 were used (Q_{427}QERLLLSVLPRHVAMEMKA$_{446}$ and Y_{1035}NRRLL HNILPKDVAA HFLA$_{1054}$, respectively; from the Uniprot entry O95622).

Plasmid construction

Standard molecular biology methods were used for DNA manipulations (a list of primers is available as Table S2). DNA fragments and vectors were restricted at their 5′-ends by *Bam*HI or *Eco*RI and at the 3′-ends by *Hin*dIII sites and inserted into pQE80 (Δ*Xho*I; Δ*Nco*I). When necessary, silent restriction sites were introduced to facilitate cloning. Constructs carried an N-terminal MRGS-His_6-tag for detection in western blots. In the pETDuet-3 vector the first MCS was replaced by that of pQE30, thus introducing the His_6-tag. The second MCS in pETDuet-3 carried a C-terminal S-tag for western blot detection. The linked heterodimers of Rv1625c were fitted with catalytic residues mutated according to those present in mammalian AC-C1 and AC-C2 domains. The linked dimeric construct was His_6-tag-Rv1625c$_{204–443}$(N372T, R376H)-TRAAGGPP AAGGRS-Rv1625c$_{204–443}$(D256S, D300S, S301T). This linker has been successfully used in a mammalian and a mycobacterial class IIIa AC [4,17]. The fidelity of all constructs was confirmed by double-stranded DNA sequencing by GATC (Konstanz, Germany).

Protein expression

Vectors with DNA constructs were transformed into *E. coli* BL21(DE3). Strains were grown overnight in LB medium (20 g LB broth·L^{-1}) at 30 °C containing 100 μg·mL^{-1} ampicillin. A 200 mL volume of LB medium (with antibiotic) was inoculated with 5 mL of a preculture and grown at 37 °C. At an A_{600} of 0.3 the temperature was lowered to 22 °C and expression was started with 500 μM (pET vectors) or 100 μM isopropyl β-D-thiogalactopyranosid (pQE vectors) for 3–4 h. Cells were harvested by centrifugation, washed once with 50 mM Tris/HCl, 1 mM EDTA, pH 8 and stored at −80 °C. For preparation of cell membranes, cells were suspended in lysis buffer (50 mM Tris/HCl, 2 mM thioglycerol, 50 mM NaCl, pH 8) and disintegrated with a French press (1100 p.s.i.). After removal of cell debris (4300 ***g***, 30 min, 4 °C), membranes were collected at 100 000 ***g*** (1 h at 4 °C). Membranes were suspended in buffer (40 mM Tris/HCl, pH 8, 1.6 mM thioglycerol, 20% glycerol) and assayed for AC activity. For preparation of cytosolic proteins cell debris and membranes were removed (48 000 ***g***, 30 min, 4 °C) and 200 μL Ni-NTA slurry (Qiagen, Hilden, Germany) added to the supernatant. After gentle rocking for 3 h on ice, the resin was poured into a column and washed (2 mL per wash). Wash buffer A was 50 mM Tris/HCl pH 8, 10 mM β-mercaptoethanol, 2 mM $MgCl_2$, 400 mM NaCl, 5 mM imidazole; wash buffer B was 50 mM Tris/HCl pH 8, 10 mM β-mercaptoethanol, 2 mM $MgCl_2$, 400 mM NaCl, 15 mM imidazole; and wash buffer C was 50 mM Tris/HCl pH 8, 10 mM β-mercaptoethanol, 10 mM $MgCl_2$, 10 mM NaCl, 15 mM imidazole. Proteins were eluted with 0.5–1 mL of buffer C containing 150 mM imidazole. The eluates were dialyzed overnight against 50 mM Tris/HCl pH 7.5, 2 mM β-mercaptoethanol, 10 mM $MgCl_2$ and 20% glycerol and assayed for AC activity. Purified proteins were stored in dialysis buffer at 4 °C.

Adenylate cyclase assay

Adenylate cyclase activity was determined for 10 min in 100 μL at 37 °C. The reactions contained 1.5 μg protein, 50 mM Tris/HCl pH 8, 22% glycerol, 3 mM $MnCl_2$, 6 mM creatine phosphate and 230 μg creatine kinase, 75 μM [α-^{32}P] ATP, and 2 mM [2,8-^{3}H]cAMP to monitor yield during cAMP purification [4,40]. Substrate conversion was kept below 10%. Ligand was added in DMSO not exceeding 1 μL solvent per tube. Respective controls were carried out throughout.

Western blot analysis

The integrity of all expressed proteins was monitored by western blotting. Sample buffer was added to 1 μg of membrane proteins and applied to SDS/PAGE (12%). Proteins were blotted onto a poly(vinylidene difluoride) membrane, incubated with an RGS-His_4-antibody (Qiagen) or S-tag antibody (Novagen R&D Systems, Darmstadt, Germany) and subsequently with a 1 : 2500 dilution of the fluorophore conjugated secondary antibody Cy3 (ECL Plex goat-α-mouse IgG-Cy3; GE Healthcare, Chicago, IL, USA). Detection was

carried out with the Ettan DIGE Imager (GE Healthcare). Proteolysis of expressed products was not observed.

Bioinformatic methods

Sequence analysis

Sequences were taken from the Uniprot Reference Proteomes databank (release 2015_04). Nucleotide cyclases were identified in a HMMER 3 search (*E*-value cutoff 1×10^{-5}; [41] using EMBL SMART's CYCc family alignment [42]. We extracted the sequences of catalytic domains and CTEs, as segments N-terminal to the former, and analyzed them respectively using the CLANS software [43]. Sequence Logos were created using the WEBLOGO 3 server [44] based on the respective clusters from the cluster analysis. For the analysis of the taxonomic distribution of CTEs, we searched the uniprot20 database (version 2016_2; available at http://toolkit.tuebingen.mpg.de; [45]) with alignments built from CTE sequences of bacteria, vertebrate C1, and vertebrate C2 units using the highly sensitive tool HHBLITS [46].

Structure models

Structural models were created using MODELLER (version 9.16; [47]) from crystal structures of Rv1625c (pdb: 4p2f), a soluble type 10 AC from rat (pdb: 4clf), and the Bst2-Thetherin Ectodomain (pdb: 2x7a) for the S-helix. The straight CTE was modelled without a template using structural constraints for this segment.

Statistical analysis

Data are presented as means ± SEM where applicable. Student's *t* test was used for comparisons. Numbers of experiments are given in the legends of the figures. The estimated EC_{50}/IC_{50} concentrations were derived from respective concentration–response curves.

Acknowledgements

We thank Ursula Kurz for expert technical assistance. This work was supported in part by the Deutsche Forschungsgemeinschaft (SFB 766) and institutional funds from the Max-Planck-Society.

Conflict of interest

The authors declare that no conflict of interest exists.

Data Accessibility

Research data pertaining to this article is located at figshare.com: https://dx.doi.org/10.6084/m9.figshare.5001914 [Correction added after online publication on 13 Jun 2017: Data Accessibility section added].

Author contributions

JES conceived and coordinated the study and wrote the paper together with all coauthors. MZ and SB designed, performed and analyzed experiments shown in Figs 1 and 2. MZ designed, performed and analyzed experiments shown in Figs 4, 6 and 7. JB and JES designed and analyzed data in Figs 3 and 5. JB and ANL designed and analyzed data in Fig. 8. AS provided cloning assistance and analyzed data. All authors reviewed the results and approved the final version of the manuscript.

References

1 Barzu O & Danchin A (1994) Adenylyl cyclases: a heterogeneous class of ATP-utilizing enzymes. *Prog Nucleic Acid Res Mol Biol* **49**, 241–283.
2 Linder JU & Schultz JE (2003) The class III adenylyl cyclases: multi-purpose signalling modules. *Cell Signal* **15**, 1081–1089.
3 Sinha SC & Sprang SR (2006) Structures, mechanism, regulation and evolution of class III nucleotidyl cyclases. *Rev Physiol Biochem Pharmacol* **157**, 105–140.
4 Guo YL, Seebacher T, Kurz U, Linder JU & Schultz JE (2001) Adenylyl cyclase Rv1625c of *Mycobacterium tuberculosis*: a progenitor of mammalian adenylyl cyclases. *EMBO J* **20**, 3667–3675.
5 Beltz S, Bassler J & Schultz JE (2016) Regulation by the quorum sensor from *Vibrio* indicates a receptor function for the membrane anchors of adenylate cyclases. *Elife* **5**, e13098.
6 Ng WL & Bassler BL (2009) Bacterial quorum-sensing network architectures. *Annu Rev Genet* **43**, 197–222.
7 Spirig T, Tiaden A, Kiefer P, Buchrieser C, Vorholt JA & Hilbi H (2008) The *Legionella* autoinducer synthase LqsA produces an alpha-hydroxyketone signaling molecule. *J Biol Chem* **283**, 18113–18123.
8 Tiaden A & Hilbi H (2012) alpha-Hydroxyketone synthesis and sensing by *Legionella* and *Vibrio*. *Sensors* **12**, 2899–2919.
9 Krupinski J, Coussen F, Bakalyar HA, Tang WJ, Feinstein PG, Orth K, Slaughter C, Reed RR & Gilman AG (1989) Adenylyl cyclase amino acid sequence: possible channel- or transporter-like structure. *Science* **244**, 1558–1564.
10 Sunahara RK, Dessauer CW & Gilman AG (1996) Complexity and diversity of mammalian adenylyl cyclases. *Annu Rev Pharmacol Toxicol* **36**, 461–480.
11 Ng WL, Perez LJ, Wei Y, Kraml C, Semmelhack MF & Bassler BL (2011) Signal production and detection

specificity in *Vibrio* CqsA/CqsS quorum-sensing systems. *Mol Microbiol* **79**, 1407–1417.
12 Ng WL, Wei Y, Perez LJ, Cong J, Long T, Koch M, Semmelhack MF, Wingreen NS & Bassler BL (2010) Probing bacterial transmembrane histidine kinase receptor-ligand interactions with natural and synthetic molecules. *Proc Natl Acad Sci USA* **107**, 5575–5580.
13 Schultz JE & Natarajan J (2013) Regulated unfolding: a basic principle of intraprotein signaling in modular proteins. *Trends Biochem Sci* **38**, 538–545.
14 Winkler K, Schultz A & Schultz JE (2012) The s-helix determines the signal in a tsr receptor/adenylyl cyclase reporter. *J Biol Chem* **287**, 15479–15488.
15 Ma X, Beuve A & van den Akker F (2010) Crystal structure of the signaling helix coiled-coil domain of the beta1 subunit of the soluble guanylyl cyclase. *BMC Struct Biol* **10**, 2.
16 Kleinboelting S, Diaz A, Moniot S, van den Heuvel J, Weyand M, Levin LR, Buck J & Steegborn C (2014) Crystal structures of human soluble adenylyl cyclase reveal mechanisms of catalysis and of its activation through bicarbonate. *Proc Natl Acad Sci USA* **111**, 3727–3732.
17 Tang WJ & Gilman AG (1995) Construction of a soluble adenylyl cyclase activated by Gs alpha and forskolin. *Science* **268**, 1769–1772.
18 Tesmer JJ, Sunahara RK, Gilman AG & Sprang SR (1997) Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS. *Science* **278**, 1907–1916.
19 Yan SZ, Hahn D, Huang ZH & Tang WJ (1996) Two cytoplasmic domains of mammalian adenylyl cyclase form a Gs alpha- and forskolin-activated enzyme in vitro. *J Biol Chem* **271**, 10941–10945.
20 Dessauer CW & Gilman AG (1997) The catalytic mechanism of mammalian adenylyl cyclase. Equilibrium binding and kinetic analysis of P-site inhibition. *J Biol Chem* **272**, 27787–27795.
21 Sadana R & Dessauer CW (2009) Physiological roles for G protein-regulated adenylyl cyclase isoforms: insights from knockout and overexpression studies. *Neurosignals* **17**, 5–22.
22 Dessauer CW, Tesmer JJ, Sprang SR & Gilman AG (1998) Identification of a Gialpha binding site on type V adenylyl cyclase. *J Biol Chem* **273**, 25831–25839.
23 Anantharaman V, Balaji S & Aravind L (2006) The signaling helix: a common functional theme in diverse signaling proteins. *Biol Direct* **1**, 25.
24 Ramamurthy V, Tucker C, Wilkie SE, Daggett V, Hunt DM & Hurley JB (2001) Interactions within the coiled-coil domain of RetGC-1 guanylyl cyclase are optimized for regulation rather than for high affinity. *J Biol Chem* **276**, 26218–26229.
25 Schultz JE, Kanchan K & Ziegler M (2015) Intraprotein signal transduction by HAMP domains: a balancing act. *Int J Med Microbiol* **305**, 243–251.
26 Bhate MP, Molnar KS, Goulian M & DeGrado WF (2015) Signal transduction in histidine kinases: insights from new structures. *Structure* **23**, 981–994.
27 Dunin-Horkawicz S & Lupas AN (2010) Comprehensive analysis of HAMP domains: implications for transmembrane signal transduction. *J Mol Biol* **397**, 1156–1174.
28 Ferris HU, Dunin-Horkawicz S, Hornig N, Hulko M, Martin J, Schultz JE, Zeth K, Lupas AN & Coles M (2012) Mechanism of regulation of receptor histidine kinases. *Structure* **20**, 56–66.
29 Tews I, Findeisen F, Sinning I, Schultz A, Schultz JE & Linder JU (2005) The structure of a pH-sensing mycobacterial adenylyl cyclase holoenzyme. *Science* **308**, 1020–1023.
30 Mondejar LG, Lupas A, Schultz A & Schultz JE (2012) HAMP domain-mediated signal transduction probed with a mycobacterial adenylyl cyclase as a reporter. *J Biol Chem* **287**, 1022–1031.
31 Hulko M, Berndt F, Gruber M, Linder JU, Truffault V, Schultz A, Martin J, Schultz JE, Lupas AN & Coles M (2006) The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell* **126**, 929–940.
32 Sinha SC, Wetterer M, Sprang SR, Schultz JE & Linder JU (2005) Origin of asymmetry in adenylyl cyclases: structures of *Mycobacterium tuberculosis* Rv1900c. *EMBO J* **24**, 663–673.
33 Steegborn C, Litvin TN, Hess KC, Capper AB, Taussig R, Buck J, Levin LR & Wu H (2005) A novel mechanism for adenylyl cyclase inhibition from the crystal structure of its complex with catechol estrogen. *J Biol Chem* **280**, 31754–31759.
34 Dessauer CW & Gilman AG (1996) Purification and characterization of a soluble form of mammalian adenylyl cyclase. *J Biol Chem* **271**, 16967–16974.
35 Chen YZ, Matsushita MM, Robertson P, Rieder M, Girirajan S, Antonacci F, Lipe H, Eichler EE, Nickerson DA, Bird TD *et al.* (2012) Autosomal dominant familial dyskinesia and facial myokymia: single exome sequencing identifies a mutation in adenylyl cyclase 5. *Arch Neurol* **69**, 630–635.
36 Chen YZ, Friedman JR, Chen DH, Chan GC, Bloss CS, Hisama FM, Topol SE, Carson AR, Pham PH, Bonkowski ES *et al.* (2014) Gain-of-function ADCY5 mutations in familial dyskinesia with facial myokymia. *Ann Neurol* **75**, 542–549.
37 Chen DH, Meneret A, Friedman JR, Korvatska O, Gad A, Bonkowski ES, Stessman HA, Doummar D, Mignot C, Anheim M *et al.* (2015) ADCY5-related dyskinesia: broader spectrum and genotype-phenotype correlations. *Neurology* **85**, 2026–2035.

38 Wilkie SE, Newbold RJ, Deery E, Walker CE, Stinton I, Ramamurthy V, Hurley JB, Bhattacharya SS, Warren MJ & Hunt DM (2000) Functional characterization of missense mutations at codon 838 in retinal guanylate cyclase correlates with disease severity in patients with autosomal dominant cone-rod dystrophy. *Hum Mol Genet* **9**, 3065–3073.
39 Payne AM, Morris AG, Downes SM, Johnson S, Bird AC, Moore AT, Bhattacharya SS & Hunt DM (2001) Clustering and frequency of mutations in the retinal guanylate cyclase (GUCY2D) gene in patients with dominant cone-rod dystrophies. *J Med Genet* **38**, 611–614.
40 Salomon Y (1979) Adenylate cyclase assay. *Adv Cyclic Nucleotide Res* **10**, 35–55.
41 Eddy SR (2011) Accelerated profile HMM searches. *PLoS Comput Biol* **7**, e1002195.
42 Schultz J, Milpetz F, Bork P & Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci USA* **95**, 5857–5864.
43 Frickey T & Lupas A (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20**, 3702–3704.
44 Crooks GE, Hon G, Chandonia JM & Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* **14**, 1188–1190.
45 Alva V, Nam SZ, Soding J & Lupas AN (2016) The MPI bioinformatics toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* **44**, W410–W415.
46 Remmert M, Biegert A, Hauser A & Soding J (2012) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods* **9**, 173–175.
47 Webb B & Sali A (2014) Comparative protein structure modeling using MODELLER. *Curr Protoc Bioinformatics* **47**, 5.6.1–32.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. Sequences of constructs described in this work.

Table S2. List of primers used in this work.

C Manuscript Bassler et al. (2018) Cellular Signalling

Bassler, J., Schultz J. E., and Lupas A. N.

Adenylate cyclases: Receivers, transducers, and generators of signals

in: Cellular Signalling 46 (2018) pp. 135-144

Personal contribution:

Conception of the bioinformatic analysis

Acquisition of all data

Analysis and interpretation of data in figures 1 and 2 (with JES)

Analysis and interpretation of data in figures 3-5 (with ANL)

Drafting and revising the manuscript (with all authors)

Cellular Signalling 46 (2018) 135–144

Contents lists available at ScienceDirect

Cellular Signalling

journal homepage: www.elsevier.com/locate/cellsig

Review

Adenylate cyclases: Receivers, transducers, and generators of signals

Jens Bassler[a], Joachim E. Schultz[b,*], Andrei N. Lupas[a,*]

[a] Max-Planck-Institut für Entwicklungsbiologie, Abt. Proteinevolution, Max-Planck-Ring 5, 72076 Tübingen, Germany
[b] Pharmazeutisches Institut der Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

ABSTRACT

Class III adenylate cyclases (ACs) are widespread signaling proteins, which translate diverse intracellular and extracellular stimuli into a uniform intracellular signal. They are typically composed of an N-terminal array of input domains and transducers, followed C-terminally by a catalytic domain, which, as a dimer, generates the second messenger cAMP. The input domains, which receive stimuli, and the transducers, which propagate the signals, are often found in other signaling proteins. The nature of stimuli and the regulatory mechanisms of ACs have been studied experimentally in only a few cases, and even in these, important questions remain open, such as whether eukaryotic ACs regulated by G protein-coupled receptors can also receive stimuli through their own membrane domains. Here we survey the current knowledge on regulation and intramolecular signal propagation in ACs and draw comparisons to other signaling proteins. We highlight the pivotal role of a recently identified cyclase-specific transducer element located N-terminally of many AC catalytic domains, suggesting an intramolecular signaling capacity.

1. Introduction

Adenosine 3′,5′-monophosphate (cAMP) is a general-purpose signaling molecule present in most branches of life. Intracellular cAMP levels control diverse cellular functions. In bacteria, cAMP regulates metabolism by activation of the catabolite activator protein (CAP), also known as cAMP receptor protein (CRP) [1]. Other cAMP-dependent responses include phototaxis [2], protein secretion, and virulence [3]. In eukaryotes, cAMP acts mainly as a downstream messenger of G protein-coupled receptors, which are involved in many sensory and developmental processes [1,4]. cAMP-producing enzymes were traditionally termed adenylate cyclases (or, from another chemical point of view adenylyl cyclases; abbreviated as ACs; see Dessauer et al. [4] for the history of AC designations), which, despite their common function, do not all originate from a common ancestor. Rather, ACs are currently divided into six unrelated classes, five of which (classes I, II, IV, V, VI) have not been studied in great detail, mostly because they are limited to a narrow range of prokaryotic species [5]. Among them, class II ACs stand out in that they are secreted as toxins by a number of pathogenic bacteria, such as *Bacillus anthracis* or *Bordetella pertussis*, perturbing cAMP levels in their hosts [6–8]. Novel evidence for the existence of nucleotide cyclase activity in plants suggest the presence of one or more additional classes of ACs that await detailed molecular, biochemical and bioinformatic characterization [9].

Class III, finally, is the numerically largest, structurally and functionally most diverse, and pharmacologically most relevant; it is also the only one occurring in animals. The family is defined by its conserved catalytic domains. It is related to the catalytic GGDEF domain of bacterial diguanylate cyclases [10]. Their common ancestor may have evolved from an early nucleotide polymerase (indicated by dotted lines in Fig. 1; [11,12]). Class III also includes the eukaryotic guanylate cyclases (GCs), whose role in regulating cellular processes overlaps little, if at all, with that of ACs. The existence of bacterial GCs is controversial [13]. Two cases of GC activity have been reported from *Rhodospirillum* [14,15]. However, many ACs show side-activities with other nucleotide triphosphates, such as the mycobacterial Rv1900c, which shows up to 7% GC side activity [16,17].

Class III ACs require dimerization for activity. Bacterial ACs are homodimers, which form two catalytic centers at the subunit interface. Eukaryotic ACs, including all ten human isoforms, are so-called pseudoheterodimers, which consist of two complementary catalytic units joined into a single chain, forming one catalytic center at the subunit

Abbreviations: AC, adenylate cyclase; BLUF, blue light receptor using FAD domain; CACHE, calcium channel and chemotaxis receptor domain; cAMP, Adenosine 3′,5′-monophosphate; CHASE, cyclase/histidine kinase-associated sensing extracellular domain; DHp, Dimerization and histidine phosphotransfer domain of histidine kinases; HAMP, transducer domain found in histidine kinases, adenylate cyclases, methylated chemotaxis proteins, and phosphatases; mAC, membrane-bound adenylate cyclase; PAS, Per (period circadian protein)/Arnt (aryl hydrocarbon receptor nuclear translocator protein)/Sim (single-minded protein) domain; sAC, soluble adenylate cyclase; SMBD, small molecule binding domain; STAND, signal transducing ATPases with numerous domains; TarH, taxis towards aspartate receptor (tar) homology domain; TPR, tetratricopeptide motif

* Corresponding authors.
E-mail addresses: joachim.schultz@uni-tuebingen.de (J.E. Schultz), andrei.lupas@tuebingen.mpg.de (A.N. Lupas).

https://doi.org/10.1016/j.cellsig.2018.03.002
Received 7 February 2018; Received in revised form 9 March 2018; Accepted 12 March 2018
Available online 18 March 2018

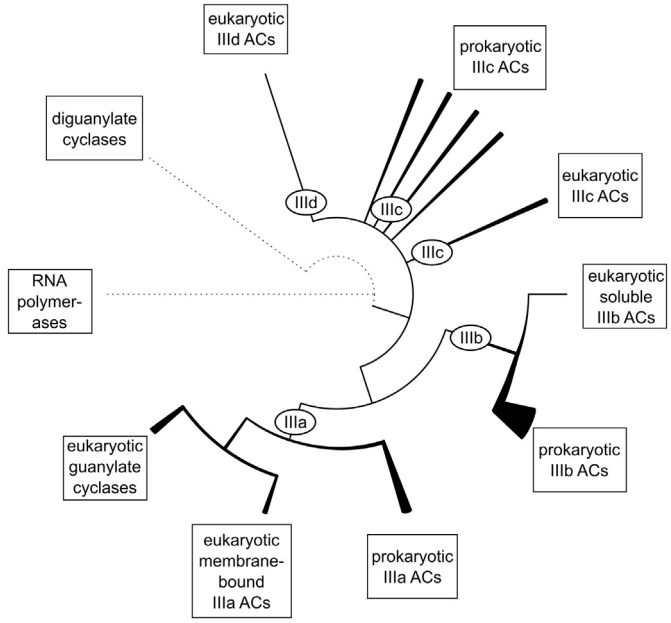


Fig. 1. The evolutionary relationship between the catalytic domains of class III adenylate cyclases. Dotted lines represent remote homology to other protein families. Solid lines represent relations between major subgroups of class III ACs. Within class III ACs, the line thickness indicates the diversity of domain architectures within a branch. Note that subclasses IIIc and IIId consist of several groups of equal rank and are thus in the text collectively referred to as subclass IIIc/d.

interface. The catalytic mechanism has been elucidated and confirmed in biochemical, structural, and computational studies [18–20]. Three pairs of residues are of particular importance: A pair of aspartate residues (Me labels in Fig. 2) coordinates a divalent metal cofactor, Mn^{2+} or Mg^{2+}, which enables a nucleophilic attack of the ribose 3′-hydroxyl group onto the ATP α-phosphoryl group. The resulting transition state is stabilized by one arginine and one asparagine side chain (Tr labels in Fig. 2). The third pair of residues discriminates between ATP and GTP as substrate (Ad labels in Fig. 2). In ACs, these are typically lysine and aspartate, whereas eukaryotic GCs frequently have glutamate and cysteine, or glutamate and serine in these positions [5].

2. Classifications of class III ACs

The widespread presence and functional diversity of class III ACs has led to several attempts at subclassification, not least in the interest of a universal ontology that could help to clarify the role of an individual AC within its cellular signaling network. The most widely accepted subclassification of class III ACs is based on sequence similarities between the homologous catalytic domains, resulting in four subclasses, termed IIIa-IIId ([5]; Figs. 1 and 2). The deepest branch-point separates IIIa and IIIb on the one hand from IIIc and IIId on the other. Within the IIIa/b branch, subclasses IIIa and IIIb are each monophyletic, i.e. represent single evolutionary lineages with characteristic features; for example, subclass IIIb ACs show an aspartate-to-threonine substitution in a substrate-determining position, which is possibly linked to bicarbonate sensitivity (Fig. 2; [21]). Subclass IIIa further divides into several subclades, corresponding to a mostly bacterial branch that probably represents the ancestral state, two clades for the catalytic domains in the animal pseudoheterodimers, one that encompasses most eukaryotic GCs, and several minor clades from diverse protozoans [22].

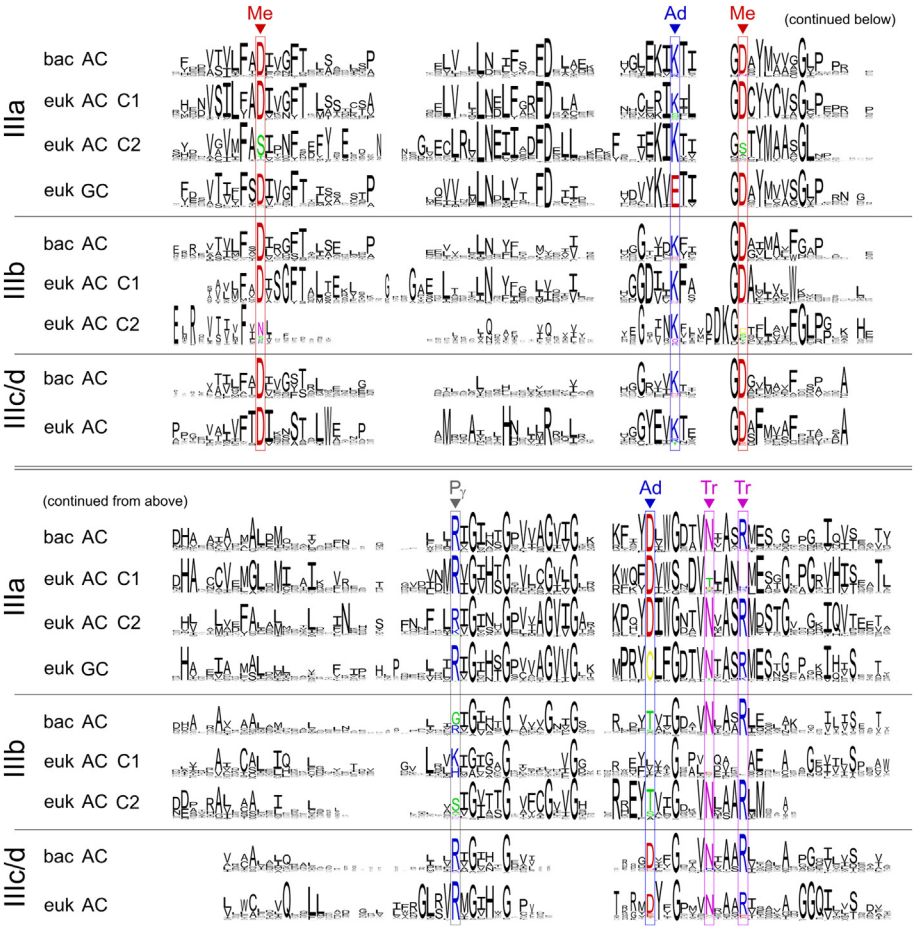


Fig. 2. Sequence logos of AC catalytic domains. The logos were created from 2387 sequences of subclass IIIa (1206 ACs, 1181 GCs), 3746 sequences of subclass IIIb, and 1823 sequences of subclass IIIc/d, clustered to 90% pairwise identity from the full dataset and plotted using WebLogo [99]. Note complementary sequences of eukaryotic C1 and C2 domains. Positions of functional importance are labeled; Me: residues coordinating the divalent metal cofactors Mg^{2+} or Mn^{2+}; Ad: residues conferring substrate specificity. In ACs, these are typically Lys/Asp (Lys/Thr in subclass IIIb). Tr: residues stabilizing the transition state; Pγ: Residue binding the γ-phosphate of the substrate.

Within the IIIc/d branch, by comparison, several clades separate near the root, of which IIId is one, the others being collected into subclass IIIc (Fig. 1; [23]). Presently, these clades have few members and do not include well-studied proteins, offering little incentive for further subclassification. In the following, we will therefore use IIIc/d as an umbrella term for the various minor lineages forming the outgroup to subclass IIIa and IIIb ACs.

In terms of the full-length proteins, it was noted early on that ACs have two cellular localizations: as integral proteins of the plasma membrane and as soluble proteins in the cytosol. Based on the distribution in mammalian sequences, it was initially proposed that subclass IIIa contains only membrane-bound forms and IIIb only soluble ones. With the increasing availability of AC sequences, it became apparent that transmembrane domains were only one aspect of AC architectural diversity [24,25] which in fact did not correlate with the IIIa-IIIb division in bacteria, where each subclass contains both forms.

Several studies have sought to catalogue the domain diversity of ACs (e.g. [26,27]) and derive from it information on general principles of regulation [28,29]. While the observed domain architectures did not lend themselves to a meaningful classification and their overall number kept growing with the number of sequenced genomes, it was found that the constituent AC domains could be grouped functionally into: (i) input domains, which receive direct stimulation (e.g. by the intracellular pH, by the binding of a ligand, by phosphorylation, or by light); (ii) transducer domains, which propagate signals from upstream input domains or other transducers; and (iii) the AC catalytic domain, which generates the second-messenger signal. Much of the observed architectural diversity can be explained by the modular recombination of input and transducer domains within this layout, arguably reflecting the physiological role of ACs as converters of a large diversity of stimuli into a uniform intracellular cAMP signal. As a general rule, input domains and transducers constitute the N-terminal part of the proteins, with the catalytic domain almost invariably located at the C-terminus. This closely parallels the arrangement in other families of signaling proteins, such as histidine kinases and diguanylate cyclases. All these proteins frequently form a coiled-coil backbone in the dimer, along which the intramolecular signal is propagated (see e.g. the structures of the light-activated AC bPAC [30], the chemoreceptor Tsr [31,32], and the histidine kinases VicK and NarQ [33,34]). Correspondingly, all input and transducer domains in these proteins are compatible with a coiled-coil structure, both for receiving and emitting conformational changes – an observation which readily explains their suitability for modular recombination into diverse architectures.

3. The phylogenetic distribution of class III ACs

To get an up-to-date view of domain architectures, we extracted class III ACs from the Uniprot reference proteomes (release 2016_11, consisting of 4058 genomes from bacteria, 229 from archaea, and 793 from eukaryotes) in an iterative HMMer3 search [35], using the catalytic domain of Rv1625c from *Mycobacterium tuberculosis* as a query (Uniprot identifier P9WQ35). Bacterial class III ACs were predicted in practically all major lineages, although frequently not in every species thereof. Only eight bacterial clades containing three or more fully sequenced species in different genera had no AC predicted: Coriobacteria and Rubrobacteria of the phylum Actinobacteria; Erysipelotrichia, Tissierellia, and Negativicutes of the Firmicutes phylum; Elusimicrobia; Synergistia; and Thermodesulfobacteria. Almost two thirds of the predicted bacterial ACs belong to subclass IIIb (3949 proteins in our dataset), followed in decreasing abundance by IIIc/d (1633 proteins, of which 422 belong to the IIId branch), and IIIa (669 proteins). While most bacteria encode only one or a few ACs, some species possess many predicted AC genes, such as *M. tuberculosis*, which encodes a total of 15 ACs from all subclasses, several of which have been studied experimentally [36,37]. Even more ACs are found in *Mycobacterium* sp. 1164966.3, with 52 AC genes, or *Turneriella parva*, with 42 AC genes, for which, however, so far only bioinformatic evidence exists.

We also observed a small number of class III ACs from archaeal species, primarily from Thaumarchaeota and Euryarchaeota. Interestingly, the archaeal ACs group closely together with bacterial ACs from subclasses IIIb and IIIc/d, and occur in similar domain architectures, suggesting that they might have been acquired through horizontal gene transfer. A more comprehensive statement about ACs in archaea is at this time not possible, due to the much lower number of sequenced species compared to bacteria and eukaryotes.

In eukaryotes, class III ACs (including GCs) are almost universally present, with the noteworthy exception of higher plants. Our dataset comprises 9690 sequences of class III nucleotide cyclases from 710 eukaryotic species. Approximately 80% thereof belong to subclass IIIa and 10% to subclasses IIIb and IIIc/d, respectively. Most sequences in subclasses IIIa and IIIb originate from animals (7153 protein sequences from 243 species), including vertebrates (3131 sequences from 101 species). Other clades in this group include Alveolates (270 sequences from 32 species), green algae (169 sequences from 10 species), fungi (79 sequences from 5 species), and a small number of brown algae and basal plants. Approximately one third of the eukaryotic subclass IIIa nucleotide cyclases can be predicted to have GTP substrate-specificity (3422 sequences from 259 species). Interestingly, the GCs from animals, green algae (Chlorophyta), diatoms, and two clades of fungi (Blastocladiomycota and Chytridiomycota) are encompassed in one homologous group, but not the ones from Amoebae and deeply branching protozoans, such as *Plasmodium*, which could have evolved their substrate-specificity independently [38,39]. Subclass IIIc/d, on the other hand, is absent from animals and found entirely in fungi, diatoms, and green algae.

As in prokaryotes, the number of ACs and GCs can vary greatly between eukaryotic species, such as in *Chlamydomonas reinhardtii* and *Naegleria gruberi*, which encode more than one hundred class III ACs each. Contrary to the situation in bacteria, however, eukaryotic ACs show comparatively low structural diversity and their domain architectures are frequently conserved throughout major clades; for example, in all animals we observe only two architectural types of ACs and three types of GCs (one of which is restricted to arthropods). Some of these have undergone an extensive lineage-specific expansion, leading to multiple genomic copies of the same type, such as the nine membrane-bound isoforms of humans (mAC; AC1-9; see Table 1). We note that these nine isoforms were numbered in the order of their cloning, without regards for potential similarities which were unrecognizable at the time. Later it was recognized that the membrane domains of ACs 1, 3 and 8 share similarities, as do those from isoforms 2, 4 and 7, and of 5 and 6. The membrane anchor domains of AC9 differ from all others. We further note that AC isoforms cloned from insects and worms have been named without considerations of potential mammalian next of kin. Table 1 gives an overview of AC descriptors commonly used in differing animals and their relationship among each other. Evidently, it would be preferable to harmonize the class IIIa nomenclature in the future.

4. Domain architectures of class III ACs

For assigning the domains of these proteins, we used HMMer3 profiles [35] of the SMART (version 6.0, [40]) and PFAM (version 31.0, [41]) databases, annotated to a significance threshold of E = 1E-5 and followed by interactive manual refinement. This identified a great variety of domain architectures, although not as many as reported earlier (e.g. [26]), arguably due to improved genome assemblies and domain definitions. In bacteria, most domain architectures appear to be conserved only between closely related species and only few architectures are substantially widespread (Figs. 3 and 4).

Table 1
Comparison of orthologous isoforms between animal model organisms (*Homo sapiens*/mammals, *Xenopus laevis*/amphibians, *Danio rerio*/bony fish, *Anopheles gambia/flies*, *Drosophila melanogaster/flies*, *Caenorhabdiitis elegans/nematodes*, *Clonorchis sinensis/flat worms*).

	H. sapiens	*X. laevis*	*D. rerio*	*A. gambiae*	*D. mel.*	*C. elegans*	*C. sinensis*
mAC	AC1	AC1	AC1a AC1b	AGAP000727	RUT	ACY4	CSIN111422
	–	–	–	AGAP007631	ACXD CG32301 CG32305		
	–	–	–	–	ACXA ACXB ACXC ACXE		
	AC5 AC6	AC5 AC6	AC5 AC6a AC6b	AGAP002998	CG43373		
	AC3	AC3	AC3a AC3b	AGAP009315	AC3		
	AC8	AC8	AC8	AGAP002262	AC78C		
	AC2 AC4	AC4	AC2a AC2b	AGAP010436	AC76E	ACY2	CSIN105540
	AC7	AC7	AC7				
	AC9	AC9	AC9	AGAP000090	AC35C	ACY1 ACY3	CSIN102237 CSIN105539 CSIN104754 CSIN100441
sAC	AC10	AC10	AC10	AGAP008683	–	–	–

Please note the discrepancies in nomenclature between mammals, and insects and Caenorhabditis ACs, respectively.

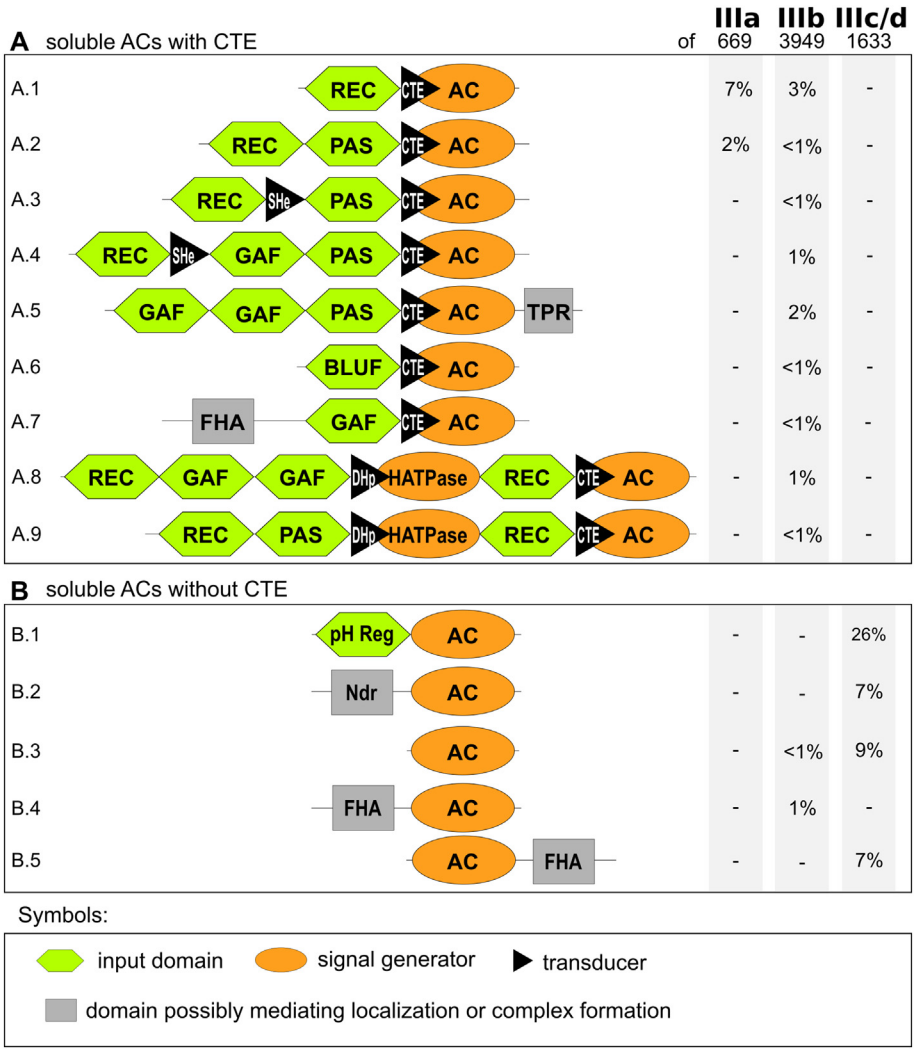


Fig. 3. Domain architectures of bacterial soluble ACs. The architectures are grouped according to the presence (panel A) or absence (panel B) of a cyclase-transducing-element (CTE). Numbers on the right indicate the abundance of this architectural type relative to the total number of bacterial sequences per subclass in our data set (IIIa: 669 sequences, IIIb: 3949 sequences, IIIc/d: 1633 sequences). Note that only predominant domain architectures are depicted, thus the numbers do not add up to 100%. Domain labels: AC, class III AC catalytic domain; CTE: cyclase transducer element; BLUF, blue light receptor using FAD domain; DHp, dimerization and histidine phosphotransfer domain of histidine kinases; FHA, forkhead-associated domain; GAF, sensor domain found in cGMP-specific phosphodiesterases, ACs, and FhlA; HATPase, histidine kinase ATPase domain; Ndr, catalytically inactive α/β hydrolase domain (name derived from the Ndr protein family); PAS, Per/Arnt/Sim domain; pH Reg, Rv1264 pH sensor homology domain; REC, (phosphorylation) receiver domain; SHe, signaling helix; TPR, tetratricopeptide motif. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

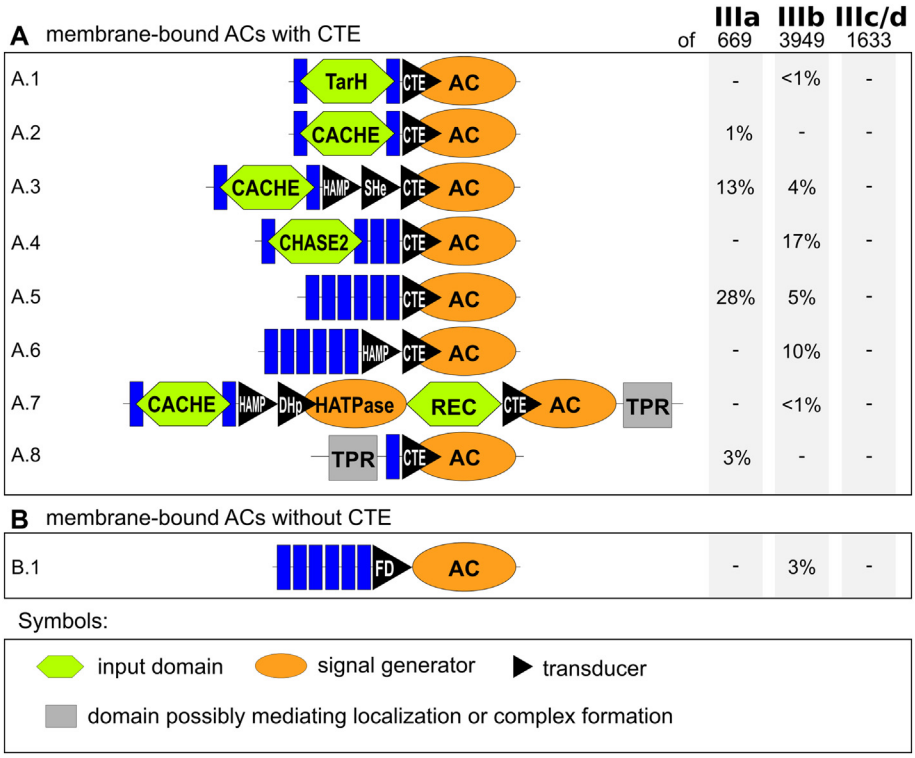


Fig. 4. Domain architectures of bacterial membrane-bound ACs. The architectures are grouped according to the presence (panel A) or absence (panel B) of a cyclase transducer element. Numbers on the right indicate the abundance of this architectural type relative to the total number of bacterial sequences per subclass in our data set (IIIa: 669 sequences, IIIb: 3949 sequences, IIIc/d: 1633 sequences). See Fig. 3 for the domain architectures of soluble bacterial ACs. Note that only selected domain architectures are displayed and the numbers thus do not add up too 100%. Blue bars indicate transmembrane helices. Domain labels: AC, class III AC catalytic domain; CACHE, calcium channel and chemotaxis receptor domain; CHASE, cyclase/histidine kinase-associated sensing extracellular domain; CTE, cyclase transducer element; DHp, dimerization and histidine phosphotransfer domain of histidine kinases; FD, ferredoxin domain; HAMP, transducer domain found in histidine kinases, adenylate cyclases, methylated chemotaxis proteins, and phosphatases; HATPase, histidine kinase ATPase domain; SHe, signaling helix; TarH, taxis towards aspartate receptor (tar) homology domain; TPR, tetratricopeptide motif. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.1. Input domains define the regulatory stimulus

A common feature shared between many ACs of subclasses IIIa and IIIb are so-called small-molecule-binding domains, such as PAS or GAF (SMBDs; [42,43]; Fig. 3(A1–7)). These are also found in other families of signaling proteins, such as histidine kinases, chemoreceptors, diguanylate cyclases, and phosphodiesterases. Generally, they act as sensors for diverse stimuli, frequently small molecules of primary metabolism, but also secondary metabolites, ions, and even photons. GAF domains, for example, have been reported in Cyanobacteria and plants to variously act as light sensors [44,45], as sodium sensors [46], or as receptors for cyclic nucleotides [47–50]. Similarly, PAS domains have been reported to have a range of ligands, including proteins and prosthetic cofactors, such as heme or FAD [51]. Distantly related to the PAS domain, BLUF domains serve as photoreceptors to a number of photoactivated ACs, such as bPAC of *Beggiatoa* sp. (Fig. 3(A.6)). Here, blue light stimulation of the FAD cofactor results in an up to 300-fold increase of AC activity [52,53], making bPAC one of the few bacterial ACs for which an experimentally tractable, physiological stimulus is known. Recently, crystal structures of bPAC and the homologous OaPAC of *Oscillatoria acuminata* partly revealed the structural rearrangement that the protein undergoes upon light activation [30,54] (see below).

In soluble ACs of subclasses IIIa and IIIb, PAS and GAF domains frequently occur in arrays (Fig. 3), indicating that they may not only act as primary sensors of stimuli, but also in stimulus propagation [47,49,55,56]. The modular shuffling of these domains in ACs has led to a sizable number of architectural variants, without a single dominant type (Fig. 3A). The genome of *Anabaena* sp., for example, encodes a total of five ACs [57], four of which belong to class III and exhibit diverse architectures (for CyaB1 and CyaB2 see Fig. 3(A.5); for CyaC Fig. 3(A.8), and for CyaD Fig. 3(A.7)).

The sensory N-terminal part of CyaB1 consists of a tandem of GAF domains, followed by a PAS domain (Fig. 3(A.5)). The tandem GAF domains activate the catalytic domain upon cAMP-binding, thus acting as autoinducers [47,55]. It has been speculated that in *Anabaena*, CyaB1 serves to stabilize committed developmental decisions [47]. A function for the C-terminal TPR-like motif is unknown.

CyaC (Fig. 3(A.8)) starts with another type of input domain, the receiver (REC) domain. REC domains frequently co-occur with SMBD domains, but, unlike these, do not bind small molecules, rather serving as phosphorylation sites for histidine kinases [58]. As such, REC domains found at the N-terminus of ACs most likely act as input sites for upstream histidine kinases in extended signaling cascades. ACs may also contain internal REC domains, which are always preceded by a histidine kinase module, comprising a dimerization and histidine phosphotransfer (DHp) domain coupled to a histidine kinase ATPase domain. CyaC contains both types of REC domains, N-terminal and internal (Fig. 3(A.8)). In CyaC of *Arthrospira platensis*, the internal histidine kinase module follows the phosphorelay of canonical histidine kinases, with the γ-phosphate of ATP being transferred by the kinase domain to the DHp domain, and from there on to the internal REC domain [59]. The N-terminal REC domain is not phosphorylated in the process, indicating that it is probably targeted by a separate kinase. Similar internal histidine kinase modules also occur in combination with other N-terminal input domains (e.g. Fig. 3(A.9) and Fig. 4(A.7)), suggesting that these may have originated from the fusion of an upstream two-component histidine kinase to a phosphorylation-activated AC, similar to type A.1 of Fig. 3.

SMBD's and REC domains are absent in ACs of subclass IIIc/d. There, the only domain with a known input function is a pH-sensing domain present in proteins of subclass IIIc (Fig. 3 B.1). Physiologically, cyclases with this domain are involved in cellular pH homeostasis [60]. The domain adopts a multi-helical fold that changes its conformation upon exposure to acidic pH. As seen in crystal structures of Rv1264 from *M. tuberculosis*, at neutral pH a long α-helix at its C-terminus holds the following catalytic domains apart in an inactive conformation. Under acidic conditions, this helix collapses, allowing the catalytic domains to form a productive homodimer [61].

Other than this, proteins of subclass IIIc/d show a diversity of domains whose role as sensors in signal transduction has not yet been elucidated and which may instead, at least in part, affect cellular localization and formation of larger signaling complexes. A prominent example is the forkhead-associated (FHA) domain, which is present also in other signaling proteins, such as protein kinases and phosphatases, as

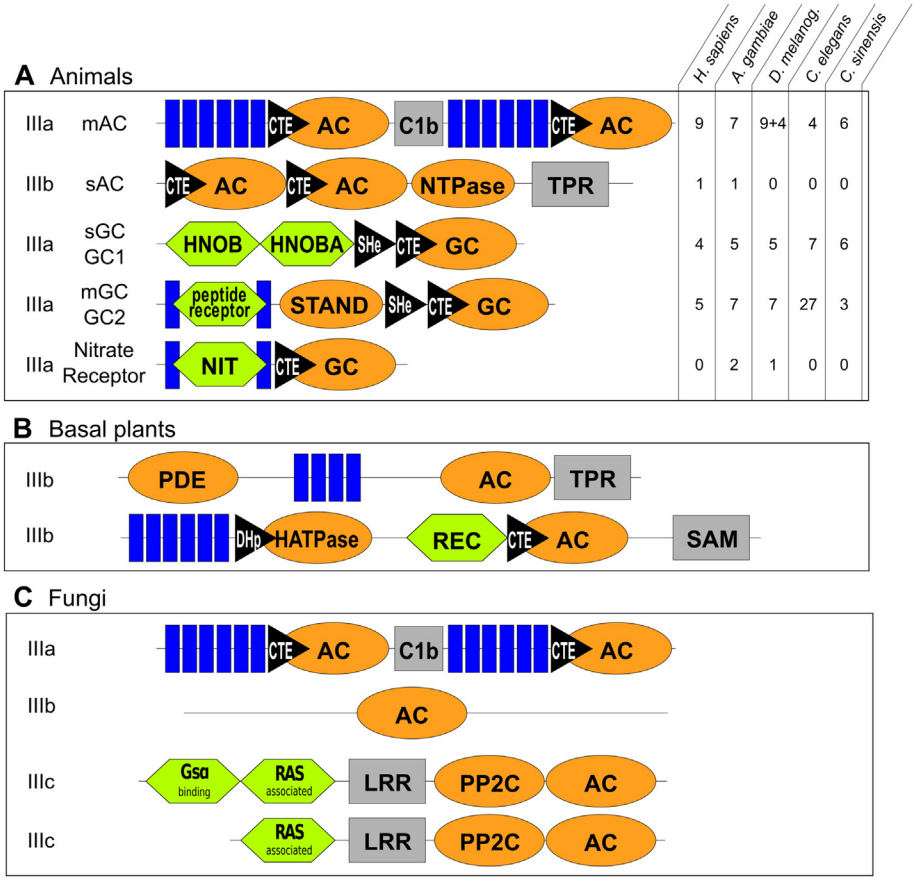


Fig. 5. Domain architectures of eukaryotic ACs. (A) Class III ACs of animals. The table on the right indicates genomic copy numbers of this architectural type in *Homo sapiens* (representing mammals), *Anopheles gambiae* (insects other than Drosophilidae), *Drosophila melanogaster* (Drosophilidae), *Caenorhabditis elegans* (Nematoda), and *Clonorchis sinensis* (Platyhelminthes). *D. melanogaster* contains four male-gonosomal mAC variants in addition to its nine autosomal forms. (B) AC domain architectures of basal plants, as observed in *Marchantia polymorpha*, *Selaginella moellendorffii*, and *Klebsormidium flaccidum*. Note that both domain architectures are also found in green algae, together with a great number of ACs with highly unusal sequences for which no other domains could be annotated. (C) AC domain architectures of Fungi. Blue bars indicate transmembrane helices. Domain labels: AC, class III AC catalytic domain; GC, class III GC catalytic domain; C1b, conserved mAC linker; CTE, cyclase transducer element; Gsα binding, unspecific Gsα binding homology domain; HNOB and HNOBA, domains mediating Heme-NO binding; HATPase, histidine kinase ATPase domain; LRR, leucin-rich repeat; NIT, Nitrate receptor; NTPase, unspecified nucleotide triphosphatase; Peptide-binding, Domain binding various peptides; PP2C, Protein phosphatase 2C homology domain; RAS associated, RAS-interacting domain; SHe, signaling helix; SAM, sterile alpha motif; STAND, signal transducing ATPases with numerous domains; TPR, tetratricopeptide motif. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

well as in proteins of unrelated function, such as kinesins, transcription factors, and metabolic enzymes [62,63]. Where known, the FHA domain is involved in binding phosphopeptides and this is thought to mediate complex assembly [64]. To what extent such a role is also able to generate an intramolecular signal through conformational change remains unknown. We note that, unlike domains with known input function, FHA and TPR (another domain frequently involved in complex formation) are found both N-terminally and C-terminally to the AC catalytic domain (e.g. Fig. 3 B.4 and B.5; Fig. 4(A.7) and (A.8)).

Where the ancillary domains of subclass IIIc/d ACs belong to enzyme families, these often carry active-site mutations and are thus likely to be catalytically inactive, substantiating a role in binding other molecules. An example is the mycobacterial Rv1900c (Fig. 3 B.2), whose N-terminal domain has an α/β-hydrolase fold (Ndr) with putative lipase activity (hence the alternative name LipJ); no such activity could be shown experimentally, pointing to a different role within the protein [16]. There are however also examples of ACs, in which these enzyme-derived domains are catalytically active, as for example a phosphodiesterase domain found in ACs of green algae and basal plants [65]; Fig. 5).

4.2. Transmembrane domains of ACs as likely input domains for extracellular stimuli

A great number of ACs in subclasses IIIa and IIIb contain one or more predicted transmembrane segments. These consist generally of two (2TM), four (4TM), or six (6TM) membrane-spanning helices (Fig. 4). The 2TM and 4TM domains possess sizable extracellular (periplasmic in Gram-negative bacteria) loops between the first and second transmembrane spans, which are typically recognized as CACHE or CHASE3/TarH domains in 2TM configurations (Fig. 4(A.1–3); [66,67]) and as CHASE2 domains in 4TM proteins (Fig. 4(A.4); [68,69]). CACHE domains likewise belong to the SMBD superfamily and have been reported as extracellular sensors in other families of signaling proteins. Known ligands in histidine kinases and chemoreceptors include various carboxylates and bivalent cations [70]. Similarly, TarH was found to sense serine or aspartate in chemoreceptors from *Escherichia coli* [71,72]. In ACs, however, no ligands have been identified so far.

By analogy to their 2TM and 4TM counterparts, 6TM domains have been considered putative input domains as well [73]. However, their constituent helices are connected by very short loops, which do not harbor any known sensory modules [74]. Thus, if 6TM domains are sensors, ligands most likely have to bind within the membrane space. Although direct evidence for this role is lacking at present, in the analogous case of 6TM domains from quorum-sensing histidine kinases, lipid-derived ligands have been identified that act as regulators of kinase activity [75–77]. Chimeric proteins in which the 6TM domain of the mycobacterial IIIa cyclase Rv1625c was replaced with the 6TM domains of the quorum sensors CqsS and LqsS showed regulation of the cyclase activity in response to their cognate ligands, Cholera-autoinducer 1 (CAI-1) and Legionella-autoindicer-1 (LAI-1), respectively [74,78], demonstrating that the cyclase dimer is fundamentally able to decode intramolecular signals.

There is currently no statistical support for a homologous relationship between the 6TM domains of histidine kinases and ACs, and even within ACs we found that the 6TM domains cluster into five distinct groups of unclear evolutionary relationship [74]. The five groups correspond to the 6TM domains of bacterial IIIa ACs (Fig. 4(A.5)); eukaryotic IIIa ACs (Fig. 5A); bacterial IIIb ACs containing a HAMP domain (Fig. 4(A.6)); bacterial IIIb ACs without HAMP (Fig. 4(A.5)); and bacterial IIIb ACs containing a ferredoxin domain (Fig. 4 B.1). The group associated with bacterial IIIa ACs (e.g. the mycobacterial Rv1625c) may be of particular relevance as a model system for the

transmembrane domains of eukaryotic IIIa ACs [79].

Several aspects are noteworthy with respect to these groups: All eukaryotic ACs with transmembrane domains are found in subclass IIIa, where the two 6TM domains of pseudoheterodimers form closely related subgroups [74]. The 6TM domains of bacterial IIIa ACs form a single group, whereas those of IIIb ACs form three clearly separated groups. Of these, the one associated with a ferredoxin domain (Fig. 4 B.1; e.g. in RPD_3704 from *Rhodopseudomonas pallustris*) does not show even remote similarity to the others, instead being clearly derived from fumarate reductases/succinate dehydrogenases [74]. Similarly, the group associated with IIIb ACs devoid of HAMP domains, such as CyaB from *Pseudomonas aeruginosa*, also shows sequence similarity to proteins external to ACs, in this case to the membrane-associated sensor (MASE2) domains of diguanylate cyclases, which have been suggested as membrane-bound input domains for an as yet unknown stimulus [80].

In summary, the 6TM domains of bacterial and eukaryotic ACs form five groups, whose sequence similarity is so distant that presently homology cannot be inferred [74]. Nevertheless, circumstantial evidence points to a role of some, possibly all, of these groups as input domains for stimuli originating outside the cell or within the membrane. In addition, within individual groups a conspicuous conservation of sequences exists, even across large evolutionary distances, which suggests that they are more than simple membrane anchors. This notion is further supported by the observation that, where the transmembrane domains of ACs have two or four helices, a sensory input function is clearly apparent. In agreement with this hypothesis, the 6TM domains can be replaced by a 2TM input module from a chemoreceptor [73,81,82] or even by an analogous 6TM from a histidine kinase [74,78], thereby regulating cyclase activity through heterologous ligands. The ability of the cyclase domain to decode intramolecular signals is also supported by the presence of HAMP domains in a sizable number of ACs with 6TM domains, as these are well-understood transducer domains (see below).

4.3. Transducers modulate signals for the catalytic domain

Beside input domains, which we define as sites of external stimulation, we observe a number of other modular elements in ACs that appear to be involved in relaying stimuli internally but are unlikely to sense external stimuli themselves. Until recently, only two such transducers were recognized in ACs: the HAMP domain and the signaling helix (S-helix).

HAMP domains are four-helical coiled-coil bundles of about 50 residues; originally named for their occurrence in histidine kinases, ACs, methylated chemotaxis proteins (chemoreceptors), and phosphatases. They are also found in diguanylate cyclases, phosphodiesterases, serine/threonine/tyrosine kinases, and other signaling proteins [23]. In membrane receptors, HAMP domains are always located immediately after the last transmembrane helix, between the upstream input domains and the downstream output domains [23]. Mutations in this domain are known to impair signaling, including locking the proteins in permanent on or off states, or even reversing the output [83–85]. In ACs, HAMP domains occur exclusively in membrane-bound proteins and strictly in single copy. Poly-HAMP arrays, as known from histidine kinases and chemoreceptors, are absent [23]. All membrane-associated AC architectures exist in configurations with and without a HAMP domain (see e.g. Fig. 4(A.2/3), Fig. 4(A.5/6)), suggesting that HAMP is not essential for decoding the transmembrane signal. We note, however, a certain mechanistic dependence between HAMP and the preceding transmembrane domain in proteins where they have co-evolved, as chimeras are more likely to be functional if HAMP and the transmembrane domains are derived from the same protein [73,81,82,86].

The second transducer found in ACs, the S-helix, forms a two-helical parallel coiled coil of typically 22 to 50 residues. As defined by Anantharaman et al. [87], the S-helix comprises a subset of dimeric coiled-coil connectors, which are widely present throughout signaling proteins. In cyclase domain architectures, we observe S-helices at different locations (Fig. 4(A.3); Fig. 3(A.3) and (A.4); Fig. 5A). A defining feature of S-helices is a central Arg-Thr motif [87,88], around which the coiled coil deviates from the canonical heptad pattern to accommodate a so-called stutter of four additional residues [89]. Deletion of the S-helix from CyaG led to an inversion of the signal, i.e. the effect of ligand binding switched from stimulation to inhibition of AC activity [78,82]. While the structural basis for this switch has not been determined in CyaG, high-resolution studies of the cognate motif in the coiled coil connecting the HAMP and DHp domains of the histidine kinase EnvZ [32] highlighted it as the part of the structure undergoing the largest transition between conformational states and showed that the main change involved an axial rotation of the coiled-coil helices. Due to the simplicity of the motif and its functional relevance, it remains unclear whether S-helices form a homologous group or converged from different starting points within the set of dimeric coiled-coil connectors.

By virtue of their structure, both HAMP and S-helices are able to propagate the signal along the coiled-coil backbone of the dimer. High-resolution structures in different conformations have led to several hypotheses concerning the mechanism of signal propagation, such as a piston-like displacement [90–92] axial rotation [93,94], and scissor-like motion [34]. Since these structures were obtained for protein fragments, it is unclear to what extent their conformations accurately reflect the degrees of freedom available to them in the context of the full proteins; correspondingly, the mechanism continues to be controversial. Irrespective of the detailed mechanism, however, all models describe a bistable equilibrium between two states. Given bistable equilibria in all domains of a receptor and an energetic coupling between adjacent domains, the state of one will stochastically determine the state of the next; a principle that has been described by many names (e.g. [28,94,95]). The result of this is a transition in the output domains between a constrained (presumed inactive) and a released (presumed active) state. This model explains why so many functional chimeras have been obtained between different receptor families, as all domains that satisfy the criteria of bistability and energetic coupling should be fundamentally exchangeable, provided that an adaptor specific to the signal generator is present, which translates the upstream motions into a uniform conformational change understandable to the catalytic output domain.

We have recently characterized such a specific transducer element in ACs, which we termed CTE for cyclase transducer element. As the name implies, and unlike general transducers, such as HAMP or S-helix, the CTE is specific for class III cyclases and appears to convert the signal from diverse upstream domains into a conformational motion that determines the activity of the AC dimer. As such, the CTE is analogous to the DHp domain of histidine kinases [78].

4.4. The CTE is a pivotal adapter to a wide range of signaling domains

We had initially noticed the CTE as a conserved sequence at the N-termini of catalytic domains of subclasses IIIa and IIIb, where it terminates the preceding coiled coil. As this suggested a suitable point for recombination, the N-terminus of the CTE was used repeatedly in chimeric constructs connecting the upstream sensory segments of various signaling proteins to the AC catalytic domain [74,78,82]. A detailed analysis of linkage points between the quorum-sensing receptor LqsS from *L. pneumophila* and the mycobacterial class IIIa AC Rv1625c determined that, for functionality, coupling had to occur at a specific sequence position, which thus defined it as the N-terminus of the CTE [78]. Connecting the domains upstream or downstream of this particular position, up to the full deletion of the conserved segment, abrogated signaling in such chimeras. Similarly, mutations in this segment had a disabling effect on intramolecular signaling; such mutations have been associated with several hereditary diseases [96–98]. This suggests that the CTE has a direct role in the regulation of AC activity. As

isolated bacterial AC catalytic domains show considerable pairwise affinity for each other, the role of the CTE probably involves the control of conformational flexibility that is required for the formation of the dimeric catalytic core [78]. In crystal structures, the CTE consists of two helices separated by a bend of approximately 45°. While the N-terminal helices of the dimer continue the upstream coiled coil, the C-terminal helices extend in opposite directions before transitioning into the catalytic domain. Based on the hydrophobic register of the C-terminal helix, which continues that of the N-terminal helix with only minor disturbance, we have proposed a model in which the CTE can adopt a continuous coiled-coil structure and thereby hold the catalytic domains apart [78]. While no such conformation has been observed, recent results from hydrogen-deuterium exchange experiments point to the CTEs as the site of the largest conformational change in catalytically active ACs [30].

As a key feature of this model, the CTE marks the endpoint of the upstream motions propagating along the central coiled-coil backbone and translates these changes into an AC-specific structural rearrangement that directly determines enzyme activity. In ACs that lack this backbone, the catalytic domains are also not preceded by CTEs (Figs. 3B and 4B) and the structural rearrangement needed for the regulation of AC activity must be provided by other mechanisms (see for example the pH-sensing Rv1264 described above). We note a clear analogy between the CTE and the DHp domain of histidine kinases, which invariably precedes the kinase domain and accepts an upstream stimulus along a coiled-coil backbone.

5. Implications for mammalian ACs

The observation that class III ACs are widely represented in bacteria, often in multiple copies per genome, whereas they are largely absent from archaea, suggests that this protein family originated in bacteria and spread into eukaryotes via the mitochondrial endosymbiont. This conjecture is further supported by the large architectural and domain diversity of bacterial ACs (Figs. 3 and 4), as contrasted with the very limited number of architectures and constituent domains found in eukaryotes (Fig. 5). In fact, the membrane-bound ACs of animals and fungi all have a single architecture, consisting of just two domain types (6TM membrane domain and AC catalytic domain) arranged in a pseudoheterodimer as TM6-AC-TM6-AC. This architecture mirrors the most widely represented form of class IIIa ACs in bacteria, which is 6TM-AC and must dimerize for activity.

In this article, we have outlined the increasing, as yet still circumstantial evidence that the 6TM domains of bacterial ACs might operate as receptors for as yet unknown ligands. Intriguingly, several architectural arguments are also applicable to the 6TM domains of mammalian membrane-bound ACs (mACs). Currently, these ACs are considered to be regulated entirely by cytosolic effectors. All mACs are regulated by G protein-coupled receptors via Gαs binding to the catalytic domains and, with the exception of mAC9, by the plant diterpene forskolin. In addition, multiple factors have been reported that affect specific isoforms, for example G-protein subunits Gαi and Gβγ, Ca^{2+} and calmodulin, and protein kinases A and C (see [4] for a detailed description of isoform-specific regulatory effects in human mACs). This raises the question as to whether the TM domains also contribute to the regulation of AC activity. Similar to the arrangement in bacterial ACs, the 6TM domains of eukaryotic mAC pseudoheterodimers are located at the N-termini of their respective modules and are connected to the catalytic domains by CTEs (see Figs. 4(A.5) and Fig. 5). This domain arrangement, particularly the presence of CTEs, indicates that signals from an N-terminal domain may affect the activity of the mACs. While the CTEs could be vestigial remnants of an ancestral signal transduction relay, their lineage-specific conservation indicates ongoing selection and a continuing functional co-evolution with the catalytic domains [78].

A similar lineage-specific conservation exists between the transmembrane domains (TM1 and TM2) of mammalian ACs as well. While the transmembrane domains mark the location of highest sequence diverge between isoforms, the sequences of the same isoform between different species are well conserved (for example, the transmembrane domains of human and mouse mACs show on average 87% sequence identity between orthologous isoforms, but only 23% identity between paralogous isoforms, compared to 98% and 56% for the respective catalytic domains C1 and C2). We take this as evidence for the existence of isoform-specific ligands. While such a function of the 6TM domains may have been inherited from a prokaryotic ancestor, in mammals regulation was expanded to G protein-dependent regulation. Because neither physiological ligands nor their binding sites are known for any membrane-delimited class III AC, and no structure of a relevant 6TM domain is available or can reliably be modelled, investigations into the hypothesized receptor functions poses a major, most interesting challenge for future research.

Acknowledgement

This work was funded by the Deutsche Forschungsgemeinschaft (SFB 766) and by institutional funds from the Max Planck Society.

References

[1] J.M. Gancedo, Biological roles of cAMP: variations on a theme in the different kingdoms of life, Biol. Rev. Camb. Philos. Soc. 88 (3) (2013) 645–668.
[2] M. Iseki, S. Matsunaga, A. Murakami, K. Ohno, K. Shiga, K. Yoshida, M. Sugai, T. Takahashi, T. Hori, M. Watanabe, A blue-light-activated adenylyl cyclase mediates photoavoidance in Euglena gracilis, Nature 415 (6875) (2002) 1047–1051.
[3] K.A. McDonough, A. Rodriguez, The myriad roles of cyclic AMP in microbial pathogens: from signal to sword, Nat. Rev. Microbiol. 10 (1) (2011) 27–38.
[4] C.W. Dessauer, V.J. Watts, R.S. Ostrom, M. Conti, S. Dove, R. Seifert, International union of basic and clinical pharmacology. CI. structures and small molecule modulators of mammalian adenylyl cyclases, Pharmacol. Rev. 69 (2) (2017) 93–139.
[5] J.U. Linder, J.E. Schultz, The class III adenylyl cyclases: multi-purpose signalling modules, Cell. Signal. 15 (2003) 1081–1089.
[6] O. Barzu, A. Danchin, Adenylyl cyclases: a heterogeneous class of ATP-utilizing enzymes, Prog. Nucleic Acid Res. Mol. Biol. 49 (1994) 241–283.
[7] S.R. Paccani, F. Finetti, M. Davi, L. Patrussi, M.M. D'Elios, D. Ladant, C.T. Baldari, The Bordetella pertussis adenylate cyclase toxin binds to T cells via LFA-1 and induces its disengagement from the immune synapse, J. Exp. Med. 208 (6) (2011) 1317–1330.
[8] A. Rogel, J.E. Schultz, R.M. Brownlie, J.G. Coote, R. Parton, E. Hanski, Bordetella pertussis adenylate cyclase: purification and characterization of the toxic form of the enzyme, EMBO J. 8 (9) (1989) 2755–2760.
[9] C. Gehring, I.S. Turek, Cyclic nucleotide monophosphates and their cyclases in plant signaling, Front. Plant Sci. 8 (2017) 1704.
[10] J. Pei, N.V. Grishin, GGDEF domain is homologous to adenylyl cyclase, Proteins 42 (2) (2001) 210–216.
[11] P.J. Artymiuk, A.R. Poirrette, D.W. Rice, P. Willett, A polymerase I palm in adenylyl cyclase? Nature 388 (6637) (1997) 33–34.
[12] J.J. Tesmer, R.K. Sunahara, A.G. Gilman, S.R. Sprang, Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS, Science 278 (5345) (1997) 1907–1916.
[13] J.U. Linder, cGMP production in bacteria, Mol. Cell. Biochem. 334 (1–2) (2010) 215–219.
[14] J.N. Marden, Q. Dong, S. Roychowdhury, J.E. Berleman, C.E. Bauer, Cyclic GMP controls Rhodospirillum centenum cyst development, Mol. Microbiol. 79 (3) (2011) 600–615.
[15] M.H. Ryu, H. Youn, I.H. Kang, M. Gomelsky, Identification of bacterial guanylate cyclases, Proteins 83 (5) (2015) 799–804.
[16] S.C. Sinha, M. Wetterer, S.R. Sprang, J.E. Schultz, J.U. Linder, Origin of asymmetry in adenylyl cyclases: structures of *Mycobacterium tuberculosis* Rv1900c, EMBO J. 24 (4) (2005) 663–673.
[17] R. Seifert, cCMP and cUMP: emerging second messengers, Trends Biochem. Sci. 40 (1) (2015) 8–15.
[18] D.K. Hahn, J.R. Tusell, S.R. Sprang, X. Chu, Catalytic mechanism of mammalian adenylyl cyclase: a computational investigation, Biochemistry 54 (40) (2015) 6252–6262.
[19] J.J. Tesmer, S.R. Sprang, The structure, catalytic mechanism and regulation of adenylyl cyclase, Curr. Opin. Struct. Biol. 8 (1998) 713–719.
[20] S.Z. Yan, Z.H. Huang, R.S. Shaw, W.J. Tang, The conserved asparagine and arginine are essential for catalysis of mammalian adenylyl cyclase, J. Biol. Chem. 272 (19) (1997) 12342–12349.
[21] M.J. Cann, A. Hammer, J. Zhou, T. Kanacher, A defined subset of adenylyl cyclases is regulated by bicarbonate ion, J. Biol. Chem. 278 (37) (2003) 35033–35038.
[22] J.H. Weber, A. Vishnyakov, K. Hambach, A. Schultz, J.E. Schultz, J.U. Linder, Adenylyl cyclases from plasmodium, paramecium and tetrahymena are novel ion

channel/enzyme fusion proteins, Cell. Signal. 16 (1) (2004) 115–125.
[23] S. Dunin-Horkawicz, A.N. Lupas, Comprehensive analysis of HAMP domains: implications for transmembrane signal transduction, J. Mol. Biol. 397 (2010) 1156–1174.
[24] A. Danchin, Phylogeny of adenylyl cyclases, Adv. Second Messenger Phosphoprotein Res. 27 (1993) 109–162.
[25] W.J. Tang, A.G. Gilman, Adenylyl cyclases, Cell 70 (6) (1992) 869–872.
[26] A.R. Shenoy, S. Visweswariah, Class III nucleotide cyclases in bacteria and archaebacteria: lineage-specific expansion of adenylyl cyclases and a dearth of guanylyl cyclases, FEBS Lett. 561 (2004) 11–21.
[27] J.U. Linder, J.E. Schultz, Versatility of signal transduction encoded in dimeric adenylyl cyclases, Curr. Opin. Struct. Biol. 18 (2008) 667–672.
[28] J.E. Schultz, J. Natarajan, Regulated unfolding: a basic principle of intraprotein signaling in modular proteins, Trends Biochem. Sci. 38 (11) (2013) 538–545.
[29] C. Steegborn, Structure, mechanism, and regulation of soluble adenylyl cyclases - similarities and differences to transmembrane adenylyl cyclases, Biochim. Biophys. Acta 1842 (12 Pt B) (2014) 2535–2547.
[30] R. Lindner, E. Hartmann, M. Tarnawski, A. Winkler, D. Frey, J. Reinstein, A. Meinhart, I. Schlichting, Photoactivation mechanism of a bacterial light-regulated adenylyl cyclase, J. Mol. Biol. 429 (9) (2017) 1336–1351.
[31] H. Tajima, K. Imada, M. Sakuma, F. Hattori, T. Nara, N. Kamo, M. Homma, I. Kawagishi, Ligand specificity determined by differentially arranged common ligand-binding residues in bacterial amino acid chemoreceptors Tsr and Tar, J. Biol. Chem. 286 (49) (2011) 42200–42210.
[32] H.U. Ferris, S. Dunin-Horkawicz, N. Hornig, M. Hulko, J. Martin, J.E. Schultz, K. Zeth, A.N. Lupas, M. Coles, Mechanism of regulation of receptor histidine kinases, Structure 20 (1) (2012) 56–66.
[33] C. Wang, J. Sang, J. Wang, M. Su, J.S. Downey, Q. Wu, S. Wang, Y. Cai, X. Xu, J. Wu, D.B. Senadheera, D.G. Cvitkovitch, L. Chen, S.D. Goodman, A. Han, Mechanistic insights revealed by the crystal structure of a histidine kinase with signal transducer and sensor domains, PLoS Biol. 11 (2) (2013) e1001493.
[34] I. Gushchin, I. Melnikov, V. Polovinkin, A. Ishchenko, A. Yuzhakova, P. Buslaev, G. Bourenkov, S. Grudinin, E. Round, T. Balandin, V. Borshchevskiy, D. Willbold, G. Leonard, G. Buldt, A. Popov, V. Gordeliy, Mechanism of transmembrane signaling by sensor histidine kinases, Science 356 (6342) (2017) 1043.
[35] S.R. Eddy, Accelerated profile HMM searches, PLoS Comput. Biol. 7 (10) (2011) e1002195.
[36] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M.A. Quail, M.A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J.E. Sulston, K. Taylor, S. Whitehead, B.G. Barrell, Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence, Nature 393 (6685) (1998) 537–544.
[37] L.A. McCue, K.A. McDonough, C.E. Lawrence, Functional classification of cNMP-binding proteins and nucleotide cyclases with implications for novel regulatory pathways in *Mycobacterium tuberculosis*, Genome Res. 10 (2) (2000) 204–219.
[38] J. Roelofs, H. Snippe, R.G. Kleineidam, P.J. Van Haastert, Guanylate cyclase in Dictyostelium discoideum with the topology of mammalian adenylate cyclase, Biochem. J. 354 (Pt 3) (2001) 697–706.
[39] D.A. Baker, Adenylyl and guanylyl cyclases from the malaria parasite *Plasmodium falciparum*, IUBMB Life 56 (9) (2004) 535–540.
[40] I. Letunic, T. Doerks, P. Bork, SMART: recent updates, new developments and status in 2015, Nucleic Acids Res. 43 (Database issue) (2015) D257–60.
[41] R.D. Finn, P. Coggill, R.Y. Eberhardt, S.R. Eddy, J. Mistry, A.L. Mitchell, S.C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G.A. Salazar, J. Tate, A. Bateman, The Pfam protein families database: towards a more sustainable future, Nucleic Acids Res. 44 (Database issue) (2016) D279–85.
[42] V. Anantharaman, E.V. Koonin, L. Aravind, Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains, J. Mol. Biol. 307 (5) (2001) 1271–1292.
[43] M.H. Hefti, K.-J. Francoijs, S.C. de Vries, R. Dixon, J. Vervoort, The PAS fold. A redefinition of the PAS domain based upon structural prediction, Eur. J. Biochem. 271 (2004) 1198–1208.
[44] B. Karniol, J.R. Wagner, J.M. Walker, R.D. Vierstra, Phylogenetic analysis of the phytochrome superfamily reveals distinct microbial subfamilies of photoreceptors, Biochem. J. 392 (Pt 1) (2005) 103–116.
[45] K. Fushimi, G. Enomoto, M. Ikeuchi, R. Narikawa, Distinctive properties of dark reversion kinetics between two red/green-type cyanobacteriochromes and their application in the photoregulation of cAMP synthesis, Photochem. Photobiol. 93 (3) (2017) 681–691.
[46] M. Cann, A subset of GAF domains are evolutionarily conserved sodium sensors, Mol. Microbiol. 64 (2) (2007) 461–472.
[47] T. Kanacher, A. Schultz, J.U. Linder, J.e. Schultz, A GAF-domain-regulated adenylyl cyclase from Anabaena is a self-activating cAMP switch, EMBO J. 21 (14) (2002) 3672–3680.
[48] J.E. Schultz, Structural and biochemical aspects of tandem GAF domains, in: H.H.H.W. Schmidt, F. Hofmann, J.-P. Stasch (Eds.), cGMP: Generators, Effectors and Therapeutic Implications, Springer-Verlag, Berlin Heidelberg, 2009, pp. 93–109.
[49] S.D. Rybalkin, I.G. Rybalkina, M. Shimizu-Albergine, X.B. Tang, J.A. Beavo, PDE5 is converted to an activated state upon cGMP binding to the GAF A domain, EMBO J. 22 (3) (2003) 469–478.
[50] R. Zoraghi, J.D. Corbin, S.H. Francis, Properties and functions of GAF domains in cyclic nucleotide phosphodiesterases and other proteins, Mol. Pharmacol. 65 (2) (2004) 267–278.
[51] J.T. Henry, S. Crosson, Ligand-binding PAS domains in a genomic, cellular, and structural context, Annu. Rev. Microbiol. 65 (2011) 261–286.
[52] M. Gomelsky, G. Klug, BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms, Trends Biochem. Sci. 27 (10) (2002) 497–500.
[53] M. Stierl, P. Stumpf, D. Udwari, R. Gueta, R. Hagedorn, A. Losi, W. Gartner, L. Petereit, M. Efetova, M. Schwarzel, T.G. Oertner, G. Nagel, P. Hegemann, Light modulation of cellular cAMP by a small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium Beggiatoa, J. Biol. Chem. 286 (2) (2011) 1181–1188.
[54] M. Ohki, K. Sugiyama, F. Kawai, H. Tanaka, Y. Nihei, S. Unzai, M. Takebe, S. Matsunaga, S. Adachi, N. Shibayama, Z. Zhou, R. Koyama, Y. Ikegaya, T. Takahashi, J.R. Tame, M. Iseki, S.Y. Park, Structural insight into photoactivation of an adenylate cyclase from a photosynthetic cyanobacterium, Proc. Natl. Acad. Sci. U. S. A. 113 (24) (2016) 6659–6664.
[55] S. Bruder, J.U. Linder, S.E. Martinez, N. Zheng, J.A. Beavo, J.E. Schultz, The cyanobacterial tandem GAF domains from the cyaB2 adenylyl cyclase signal via both cAMP-binding sites, Proc. Natl. Acad. Sci. U. S. A. 102 (8) (2005) 3088–3092.
[56] S. Bruder, A. Schultz, J.E. Schultz, Characterization of the tandem GAF domain of human phosphodiesterase 5 using a cyanobacterial adenylyl cyclase as a reporter enzyme, J. Biol. Chem. 281 (29) (2006) 19969–19976.
[57] M. Katayama, M. Ohmori, Isolation and characterization of multiple adenylate cyclase genes from the cyanobacterium Anabaena sp. strain PCC 7120, J. Bacteriol. 179 (11) (1997) 3588–3593.
[58] A.M. Stock, D.C. Wylie, J.M. Mottonen, A.N. Lupas, E.G. Ninfa, A.J. Ninfa, C.E. Schutt, J.B. Stock, Phosphoproteins involved in bacterial signal transduction, Cold Spring Harb. Symp. Quant. Biol. 53 (Pt 1) (1988) 49–57.
[59] M. Kasahara, M. Ohmori, Activation of a cyanobacterial adenylate cyclase, CyaC, by autophosphorylation and a subsequent phosphotransfer reaction, J. Biol. Chem. 274 (21) (1999) 15167–15172.
[60] U. Susstrunk, J. Pidoux, S. Taubert, A. Ullmann, C.J. Thompson, Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in Streptomyces coelicolor, Mol. Microbiol. 30 (1) (1998) 33–46.
[61] F. Findeisen Tews, I. Sinning, A. Schultz, J.E. Schultz, J.U. Linder, The structure of a pH-sensing mycobacterial adenylyl cyclase holoenzyme, Science 308 (2005) 1020–1023.
[62] K. Hofmann, P. Bucher, The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors, Trends Biochem. Sci. 20 (9) (1995) 347–349.
[63] D. Durocher, S.P. Jackson, The FHA domain, FEBS Lett. 513 (1) (2002) 58–66.
[64] A.W. Almawi, L.A. Matthews, A. Guarne, FHA domains: Phosphopeptide binding and beyond, Prog. Biophys. Mol. Biol. 127 (2017) 105–110.
[65] M. Kasahara, N. Suetsugu, Y. Urano, C. Yamamoto, M. Ohmori, Y. Takada, S. Okuda, T. Nishiyama, H. Sakayama, T. Kohchi, F. Takahashi, An adenylyl cyclase with a phosphodiesterase domain in basal plants with a motile sperm system, Sci. Rep. 6 (2016) 39232.
[66] V. Anantharaman, L. Aravind, Cache - a signaling domain common to animal ca (2+)-channel subunits and a class of prokaryotic chemotaxis receptors, Trends Biochem. Sci. 25 (11) (2000) 535–537.
[67] I.B. Zhulin, A.N. Nikolskaya, M.Y. Galperin, Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea, J. Bacteriol. 185 (1) (2003) 285–294.
[68] V. Anantharaman, L. Aravind, The CHASE domain: a predicted ligand-binding module in plant cytokinin receptors and other eukaryotic and bacterial receptors, Trends Biochem. Sci. 26 (10) (2001) 579–582.
[69] C. Mougel, I.B. Zhulin, CHASE: an extracellular sensing domain common to transmembrane receptors from prokaryotes, lower eukaryotes and plants, Trends Biochem. Sci. 26 (10) (2001) 582–584.
[70] A.A. Upadhyay, A.D. Fleetwood, O. Adebali, R.D. Finn, I.B. Zhulin, Cache domains that are homologous to, but different from PAS domains comprise the largest superfamily of extracellular sensors in prokaryotes, PLoS Comput. Biol. 12 (4) (2016) e1004862.
[71] A. Boyd, K. Kendall, M.I. Simon, Structure of the serine chemoreceptor in Escherichia coli, Nature 301 (5901) (1983) 623–626.
[72] G.L. Hazelbauer, J.J. Falke, J.S. Parkinson, Bacterial chemoreceptors: high-performance signaling in networked arrays, Trends Biochem. Sci. 33 (1) (2008) 9–19.
[73] K. Kanchan, J.U. Linder, K. Winkler, K. Hantke, A. Schultz, J.E. Schultz, Transmembrane signaling in chimeras of the Escherichia coli aspartate and serine chemotaxis receptors and bacterial class III adenylyl cyclases, J. Biol. Chem. 285 (3) (2010) 2090–2099.
[74] S. Beltz, J. Bassler, J.E. Schultz, Regulation by the quorum sensor from vibrio indicates a receptor function for the membrane anchors of adenylate cyclases, elife 5 (2016).
[75] T. Spirig, A. Tiaden, P. Kiefer, C. Buchrieser, J.A. Vorholt, H. Hilbi, The legionella autoinducer synthase LqsA produces an alpha-hydroxyketone signaling molecule, J. Biol. Chem. 283 (26) (2008) 18113–18123.
[76] W.L. Ng, B.L. Bassler, Bacterial quorum-sensing network architectures, Annu. Rev. Genet. 43 (2009) 197–222.
[77] A. Tiaden, H. Hilbi, alpha-Hydroxyketone synthesis and sensing by Legionella and Vibrio, Sensors 12 (3) (2012) 2899–2919.
[78] M. Ziegler, J. Bassler, S. Beltz, A. Schultz, A.N. Lupas, J.E. Schultz, Characterization of a novel signal transducer element intrinsic to class IIIa/b adenylate cyclases und guanylate cyclases, FEBS J. 284 (2017) 1204–1217.
[79] Y.L. Guo, T. Seebacher, U. Kurz, J.U. Linder, J.E. Schultz, Adenylyl cyclase Rv1625c of mycobacterium tuberculosis: a progenitor of mammalian adenylyl cyclases, EMBO J. 20 (14) (2001) 3667–3675.
[80] A.N. Nikolskaya, A.Y. Mulkidjanian, I.B. Beech, M.Y. Galperin, MASE1 and MASE2:

two novel integral membrane sensory domains, J. Mol. Microbiol. Biotechnol. 5 (1) (2003) 11–16.
[81] L.G. Mondejar, A. Lupas, A. Schultz, J.E. Schultz, HAMP domain-mediated signal transduction probed with a mycobacterial adenylyl cyclase as a reporter, J. Biol. Chem. 287 (2) (2012) 1022–1031.
[82] K. Winkler, A. Schultz, J.E. Schultz, The s-helix determines the signal in a tsr receptor/adenylyl cyclase reporter, J. Biol. Chem. 287 (19) (2012) 15479–15488.
[83] P. Ames, J.S. Parkinson, Transmembrane signaling by bacterial chemoreceptors: *E. coli* transducers with locked signal output, Cell 55 (5) (1988) 817–826.
[84] J.A. Appleman, V. Stewart, Mutational analysis of a conserved signal-transducing element: the HAMP linker of the Escherichia coli nitrate sensor NarX, J. Bacteriol. 185 (1) (2003) 89–97.
[85] P. Ames, Q. Zhou, J.S. Parkinson, Mutational analysis of the connector segment in the HAMP domain of Tsr, the Escherichia coli serine chemoreceptor, J. Bacteriol. 190 (20) (2008) 6676–6685.
[86] J.E. Schultz, K. Kanchan, M. Ziegler, Intraprotein signal transduction by HAMP domains: a balancing act, Int. J. Med. Microbiol. 305 (2) (2015) 243–251.
[87] V. Anantharaman, S. Balaji, L. Aravind, The signalling helix: a common functional theme in diverse signalling proteins, bio, Direct 25 (2006) 1795–1804.
[88] X. Ma, A. Beuve, F. van den Akker, Crystal structure of the signaling helix coiled-coil domain of the beta1 subunit of the soluble guanylyl cyclase, BMC Struct. Biol. 10 (2010) 2.
[89] A.N. Lupas, J. Bassler, Coiled coils - a model system for the 21st century, Trends Biochem. Sci. 42 (2) (2017) 130–140.
[90] M.V. Airola, N. Sukomon, D. Samanta, P.P. Borbat, J.H. Freed, K.J. Watts, B.R. Crane, HAMP domain conformers that propagate opposite signals in bacterial chemoreceptors, PLoS Biol. 11 (2) (2013) e1001479.
[91] M.V. Airola, K.J. Watts, A.M. Bilwes, B.R. Crane, Structure of concatenated HAMP domains provides a mechanism for signal transduction, Structure 18 (4) (2010) 436–448.
[92] K.M. Ottemann, W. Xiao, Y.K. Shin, D.E.J. Koshland, A piston model for transmembrane signaling of the aspartate receptor, Science 285 (1999) 1751–1754.
[93] M. Hulko, F. Berndt, M. Gruber, J.U. Linder, V. Truffault, A. Schultz, J. Martin, J.E. Schultz, A.N. Lupas, M. Coles, The HAMP domain structure implies helix rotation in transmembrane signaling, Cell 126 (5) (2006) 929–940.
[94] H.U. Ferris, S. Dunin-Horkawicz, L.G. Mondejar, M. Hulko, K. Hantke, J. Martin, J.E. Schultz, K. Zeth, A.N. Lupas, M. Coles, The mechanisms of HAMP-mediated signaling in transmembrane receptors, Structure 19 (3) (2011) 378–385.
[95] M.P. Bhate, K.S. Molnar, M. Goulian, W.F. DeGrado, Signal transduction in histidine kinases: insights from new structures, Structure 23 (6) (2015) 981–994.
[96] S.E. Wilkie, R.J. Newbold, E. Deery, C.E. Walker, I. Stinton, V. Ramamurthy, J.B. Hurley, S.S. Bhattacharya, M.J. Warren, D.M. Hunt, Functional characterization of missense mutations at codon 838 in retinal guanylate cyclase correlates with disease severity in patients with autosomal dominant cone-rod dystrophy, Hum. Mol. Genet. 9 (20) (2000) 3065–3073.
[97] A.M. Payne, A.G. Morris, S.M. Downes, S. Johnson, A.C. Bird, A.T. Moore, S.S. Bhattacharya, D.M. Hunt, Clustering and frequency of mutations in the retinal guanylate cyclase (GUCY2D) gene in patients with dominant cone-rod dystrophies, J. Med. Genet. 38 (9) (2001) 611–614.
[98] I. Vercellino, L. Rezabkova, V. Olieric, Y. Polyhach, T. Weinert, R.A. Kammerer, G. Jeschke, V.M. Korkhov, Role of the nucleotidyl cyclase helical domain in catalytically active dimer formation, Proc. Natl. Acad. Sci. U. S. A. 114 (46) (2017) E9821–E9828.
[99] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, Genome Res. 14 (6) (2004) 1188–1190.

Jens Baßler

Curriculum vitae

Herrenberger Straße 28
72070 Tübingen
✉ *jbassler@tue.mpg.de*

Education and Experience

2014–2018 **Ph.D.**, *Max Planck Institute for Developmental Biology*, Tübingen.

2012–2014 **M.Sc. in Biological Sciences**, *University of Konstanz*, Konstanz.
Focus on Protein Biochemistry and Bioinformatics

2009–2012 **B.Sc. in Biology**, *University of Hohenheim*, Stuttgart.
Focus on Genetics and Molecular Plant Physiology

2008–2009 **Military Service**, *German Air Force*.

1998–2008 **Abitur**, *Schelztor Gymnasium*, Esslingen a.N.

Publications

01 *Bassler J.*, Hernandez-Alvarez B., Hartmann M. D., and Lupas A. N.
A domain dictionary of Trimeric Autotransporter Adhesins
in: *International Journal of Medical Microbiology* 305 (2015) pp. 265-275

02 Hartmann M. D., Mendler C. T., *Bassler J.*, Karamichali I., Ridderbusch O., Lupas A. N., and Hernandez-Alvarez B.
α/β coiled coils
in: *eLIFE* 5 (2016)

03 Beltz S., *Bassler J.*, Schultz J. E.
Regulation by the quorum sensor from Vibrio indicates a receptor function for the membrane anchors of adenylate cyclases
in: *eLIFE* 5 (2016)

04 Lupas A. N. and *Bassler J.*
Coiled coils - A model system for the 21 century
in: *Trends in Biological Sciences* 42 (2017) pp. 130-140

05 Ziegler M., *Bassler J.*, Beltz S., Schultz A., Lupas A. N., Schultz J. E.
Characterization of a novel signal transducer element intrinsic to class IIIa/b adenylate cyclases and guanylate cyclases
in: *FEBS J* 284 (2017) pp. 1204-1217

06 Lupas A. N., *Bassler J.*, and Stanislaw Dunin-Horkawicz
The structure and topology of α-helical coiled coils
in: *Subcellular Biochemistry* 82 (2017) pp. 95-129 (book chapter)

07 *Bassler J.*, Schultz J. E., Lupas A. N.
Adenylate cyclases: Receivers, transducers, and generators of signals
in: *Cellular Signalling* 46 (2018) pp. 135-144