Title

The Study of Soluble Guanylate Cyclases from Choanoeca flexa

Author

Wu, Yang

Publication Date

2024

Peer reviewed|Thesis/dissertation

The Study of Soluble Guanylate Cyclases from *Choanoeca flexa*

By

Yang Wu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael A. Marletta, Co-Chair
Professor Nicole King, Co-Chair
Professor Michelle C. Y. Chang
Professor Sabeeha S. Merchant

Fall 2024

The Study of Soluble Guanylate Cyclases from *Choanoeca flexa*

by Yang Wu

Abstract

The Study of Soluble Guanylate Cyclases from *Choanoeca flexa*

by

Yang Wu

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Michael A. Marletta, Co-Chair

Professor Nicole King, Co-Chair

Nitric oxide (NO) is an indispensable gas signaling molecule and immune system effector in mammals. In the prototypical NO signaling pathway that mediates vasodilation in humans, the source of NO is a nitric oxide synthase (NOS), and NO is sensed by the heme-containing gas sensor soluble guanylate cyclase (sGC). Mammalian sGCs are α/β-type heterodimers activated by NO and catalyze the conversion of GTP to the second messenger cyclic GMP (cGMP). sGCs regulate critical physiological functions such as vasodilation and neuronal signaling, and misregulation of sGC leads to diseases. Therefore, sGC activity regulation has been a target of significant research effort. A major step towards elucidating the mechanism of mammalian sGC activation came about in 2018, when the cryo-electron microscopy structure of human sGC and an insect homolog of sGC from tobacco hornworm, *Manduca sexta*, were reported. This, combined with increasingly robust structural prediction techniques, greatly facilitated mechanistic studies of sGC and provided new tools to study diverse non-mammalian homologs of sGC proteins.

Besides mammals, NO signaling is important for a wide range of organisms including bacteria, algae, fungi, and many invertebrate animals. Animal pathways that are frequently regulated by NO include larval metamorphosis, flagellar movement and collective contractility, to name a few. From an evolutionary standpoint, choanoflagellates are an especially interesting organism to study NO signaling, as they are the closest living relatives of animals and may hold the key to understanding the evolution of NO signaling. An *in vivo* study described here established the presence of a NO signaling pathway in the colonial choanoflagellate *Choanoeca flexa*. *C. flexa* co-expresses NOS and sGC. NO stimulates sGC activity in *C. flexa* and drives the collective contraction behavior. cGMP produced by sGC is necessary for persistent contraction of *C. flexa* colonies. These data provided insight into a potential role of NO in early animals and evolution of NO signaling, and brought a NO-sensitive sGC, *C. flexa* sGC1 into the spotlight.

Initial characterization of *Cf* sGC1 *in vitro* demonstrated that it is a catalytically active homodimeric protein. Ligand binding properties and the ligand-induced activity profile of *Cf* sGC1 are both reminiscent of α/β-type sGC. These data prompted a more detailed biochemical characterization of *Cf* sGC1. Besides additional similarities in substrate kinetics, structural studies using small angle X-ray scattering revealed that *Cf* sGC1 is a homodimer with conformational asymmetry. This asymmetry may be connected to the 1-heme per homodimer heme stoichiometry of *Cf* sGC1, also determined in this study. Furthermore, like α/β-type sGC, a conformational extension was observed for *Cf* sGC1 during activation. Similarities between *Cf* sGC1 and α/β-type sGC suggest comparable mechanisms of activity regulation.

Like animal sGCs, *Cf* sGC1 displays a three-stage activation profile, suggesting that NO binding to heme alone is insufficient for full activation. In α/β-type sGC, cysteines were hypothesized to mediate secondary NO interaction. Involvement of cysteines in controlling the activity of *Cf* sGC1 was tested. *Cf* sGC1 treated with the cysteine labeling reagent methyl methanethiosulfonate (MMTS) was inhibited, suggesting that MMTS treatment either blocked non-heme NO interaction, or blocked conformational change at a critical cysteine site. Cysteine variants of *Cf* sGC1 at conserved sites to α/β-type sGC were prepared and characterized. Unexpectedly, cysteine variants of *Cf* sGC1 did not exhibit different properties in heme ligand binding, ligand-induced activation or MMTS-mediated inhibition. Additional work is required to definitively show the link of cysteines to activity regulation in *Cf* sGC1.

Under aerobic conditions, α/β-type sGCs and *Cf* sGC1 do not bind O_2. The study of ligand selectivity in α/β-type sGCs led to the discovery of sGC that can bind O_2. However, the mechanism for activity regulation of O_2-binding sGCs is not well understood. Besides *Cf* sGC1, the genome of *C. flexa* also encodes sGCs that bind O_2. Here, the O_2-binding sGC, *Cf* sGC4 was characterized using biochemical and structural techniques. Unlike homologs that bind NO selectively, *Cf* sGC4 is inhibited by ligand binding, including NO, CO and O_2, and does not undergo conformational change in the presence of NO. These results support a different and currently unknown mechanism of activity regulation in *Cf* sGC4 and possibly other O_2-binding sGCs.

TABLE OF CONTENTS

LIST OF ABBREVATIONS

1-NO	one equivalent nitric oxide
AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride
ASW	artificial seawater medium
ATP	adenosine triphosphate
BCA	bicinchoninic acid assay
BIC	butyl isocyanide
BLAST	Basic Local Alignment Search Tool
CaM	calmodulin
CAT	catalytic domain
CC	coiled coil domain
Cf	*Choanoeca flexa*
CGM3	cereal grass medium
CNG	cGMP-gated ion channels
CO	carbon monoxide
Cp	*Choanoeca perplexa*
Cr	*Chlamydomonas reinhardtii*
Cryo-EM	cryo-electron microscopy
Cs	*Caldanaerobacter subterraneous*
CV	column volume
DAF-FM	4-Amino-5-Methylamino-2',7'-Difluorofluorescein Diacetate
d-ALA	delta-aminolevulinic acid
DEA-NONOate	(Z)-1-(N,N-diethylamino)diazen-1-ium-1,2-diolate
Dm	*Drosophila melanogaster*
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetate
EFA	evolving factor analysis
ELISA	enzyme-linked immunosorbent assay
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
H_4B	tetrahydrobiopterin
H_4F	tetrahydrofolate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HNOX	heme nitric oxide oxygen binding domain
HPLC	high performance liquid chromatography
Hs	*Homo sapiens*
IPTG	Isopropyl β-d-1-thiogalactopyranoside
LB	Luria broth

MMTS	methyl methanethiosulfonate
MS	mass spectrometry
Ms	*Manduca sexta*
MW	molecular weight
MWCO	molecular weight cutoff
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NO	nitric oxide
NOS	nitric oxide synthase
O_2	oxygen
ODQ	oxadiazol-quinoxalin-1-one
PAS	per-arnt-sim-like domain
PDB	protein data bank
PDE	phosphodiesterase
PKG	protein kinase G
proliNONOate	1-(hydroxy-NNO-azoxy)-L-proline, disodium salt
R_g	radius of gyration
Rn	*Rattus norvegicus*
SAXS	small angle x-ray scattering
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
sGC	soluble guanylate cyclase
So	*Shewanella oneidensis*
Sy	*Synechococcus sp.* PCC7335
SSN	sequence similarity network
TB	terrific broth
TCEP	tris(2-carboxyethyl)phosphine
UV-vis	ultraviolet-visible
WT	wild type
xsNO	excess nitric oxide

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

During the past several years as I pursued my Ph.D. education, I have received the kindness and support of my mentors, colleagues, friends and family, and I would like to take this opportunity to express my gratitude.

I would like to thank my research advisor, Professor Michael Marletta. It has been a privilege to work with you. I have enjoyed learning from your vast knowledge in biochemistry, and I appreciate your patience as I learned how to do science well and how to communicate science effectively. I have enjoyed your mentorship style, and I always felt encouraged to pursue new directions of research independently. I also loved working with the team that you helped assemble. All said, I cannot ask for a better mentor than the one you have been.

To my committee co-chair Professor Nicole King, thank you for generously accepting me to be a part of your research program. The opportunity to work with a group of extraordinary molecular biologists was amazing, and I enjoyed the informative and thought-provoking discussions that took place during our meetings and meetings with members of the King Lab. I am glad to have had the opportunity to contribute my small part to understanding NO signaling in choanoflagellates.

I would also like to thank my committee members Professor Michelle Chang and Professor Sabeeha Merchant, particularly for sharing their wisdom as I wrapped up my graduate work. My achievements would not have been possible without my previous teachers and mentors. I would like to thank my undergraduate research mentor, Professor Louise Charkoudian who introduced me to the career in biochemical research, and my middle school chemistry teacher Shanlan Hu who nurtured my newfound passion in chemistry.

During the past five years I have worked with amazing colleagues: Allie Batka, Dr. Tyler Detomasi, Kimberly Houghton, Dr. Christopher Lemon, Dr. Alejandra Martinez, Dr. William Thomas, Dr. Richard Sayler, and Dr. Elizabeth Wittenborn. I have learned so much from each of you, and you have made coming into the lab something that I look forward to. To the mentees that I've had the opportunity to work with Zachary Green and Valerie Jin, I appreciate the energy that you bring to the lab. Thank you for being patient with me as I learned the ropes of being a mentor. To the researchers I collaborated with, Dr. Josean Reyes-Rivera, Dr. Thibaut Brunet, Jack Feldner, Kate Dolph, Celine Wang and Polina Barzova, thank you for the insightful discussions throughout our collaboration, and it was always fun working with you.

To the administration team at Stanley Hall, Kris Thompson, Cherry Chung, Thom Opal, at the receiving department, Mike Bently and Emily Rodriguez, our labware cleaning technician Crystal, and to all the custodial staff of Stanley Hall, thank you for all the work behind the scenes. You have made our lab space in Stanley Hall a clean, safe and efficient place to work in, and your work is a part of my every successful experiment.

Outside of science, to my friends Yutong Li, Dr. Boyu Qie, Ziyi Wang, and Dr. Zhiling Zheng, thank you. I will miss our dinner discussions on the current state of scientific research and science policies. My family has been very supportive towards me during my graduate training. To my parents, thank you for your encouragement as I pursued my passion in science. And to my siblings Richard and David, I appreciate your interest in my work, and I've always enjoyed chatting with you about science.

Chapter 1
Introduction

Overview of NO-cGMP signaling

Nitric oxide (NO) is a crucial diatomic signaling agent for bacteria and eukaryotes alike.[1] In mammals, NO at nanomolar concentrations is the signal that drives vasodilation, but at higher concentrations, NO synthesized by macrophages is used as an effector molecule in the host immune response.[2–5] Additional evidence has emerged to implicate NO in other important pathways in animal physiology, including in neurotransmission and metabolic regulation.[6–8] It was not until the 1980s when it was definitively shown that NO drives vasodilation through the NO-cGMP axis, and the work that led to this discovery was credited with the Nobel Prize in physiology and medicine in 1998. Dysregulation of NO signaling can cause pathologies in several different tissues and pathways, including neurodegeneration, inflammatory diseases and cardiovascular diseases.[9–11] In more recent years, the mammalian sensor of NO, soluble guanylate cyclase (sGC), has become a drug target for diseases like pulmonary hypertension, with an FDA-approved drug Adempas (Riociguat) approved in 2013.[12]

NO-sensing pathways can be separated into the NO source and the NO receptor. The source of NO in well characterized vasculature and neuronal signaling pathways is nitric oxide synthase (NOS).[13,14] The same is true for macrophage-derived NO. In eukaryotes, sGCs are the sensors of NO that, upon activation by NO, synthesize the second messenger cGMP to amplify the signal.[4,15] cGMP then interacts with protein kinase G, phosphodiesterases, or cGMP-dependent ion channels to drive downstream signaling.[16] In this chapter, a general introduction to the structure, activity and regulation of NO signaling proteins NOS and sGC pertaining to the projects detailed in this thesis is provided. Diverse sGC-mediated gas sensing pathways in non-mammal, non-insect animals, and single celled eukaryotes will also be highlighted.

Nitric oxide synthesis: nitric oxide synthase

The source of NO in NO-cGMP signaling in animals is typically a NOS. In mammals, NOSs are obligate homodimeric proteins, composed of an N-terminal oxidase domain and a C-terminal reductase domain.[17] The N-terminal oxidase domain acts like a heme dependent, P450-like monooxygenase. The C-terminal reductase domain utilizes redox cofactors NADPH and enzyme-bound FMN and FAD to sequentially deliver electrons to the oxidase domain to complete the catalytic cycle.[18] Additional cofactors in the mammalian NOS include tetrahydrobiopterin (H_4B) and a structural zinc, which are crucial to electron delivery and dimerization, respectively.[19–21] The reaction catalyzed by NOS is the 5-electron oxidation of L-arginine to L-citrulline via the intermediate N-hydroxy-L-arginine using O_2 as the co-substrate. In humans, three tissue-specific isoforms of NOS have been identified. The neuronal NOS (nNOS) and endothelial NOS (eNOS) are constitutively expressed in neurons and vascular endothelial cells, respectively, and their activation require Ca^{2+} release and the subsequent binding of Ca^{2+}/calmodulin.[22,23] On the other hand, the inducible NOS (iNOS) is expressed in macrophages and is transcriptionally regulated.[5] Ca^{2+}/Calmodulin is constitutively bound to iNOS and is not dependent on Ca^{2+} release.[24,25] Thus, iNOS is constitutively active.

Nitric Oxide Synthase in bacteria

NO is used as a signaling molecule by many organisms other than mammals, so NOSs are found in a wide range of organisms. Single domain NOSs were first identified in the bacterial species *Bacillus subtilis* and *Deinococcus radiodurans* and they use standalone reductases for catalysis (Figure 1.1).[26,27] Also in bacteria, there exist NOSs that are encoded with an iron sulfur cluster reductase domain in the same reading frame (Figure 1.1).[28] While the P450-like oxygenase domain of bacterial NOSs share the same overall structural fold to the oxidase domain of animal NOSs, some key differences remain, such as their lack of a zinc binding site at the dimer interface, and their dependence on tetrahydrofolate (H_4F) instead of H_4B.[29]

In cyanobacteria, NOS with a domain architecture that better resembles that of animal NOS has also been found. One such NOS was characterized in 2018, the NOS from the cyanobacteria *Synechococcus sp.* PCC 7335 (Figure 1.1).[30] *Sy* NOS is encoded within one contiguous reading frame and contains both an N-terminal oxidase domain and a C-terminal reductase domain that harbors NADPH, FAD and FMN binding sites. *In vitro* activity data suggest that *Sy* NOS uses H_4B as a cofactor. However, *Sy* NOS does not contain a Ca^{2+}/calmodulin binding site. In mammalian NOSs, binding of Ca^{2+}/calmodulin triggers a conformational change that brings the reductase domain into proximity of the oxidase domain to facilitate the transfer of electrons and enable catalysis.[18] The lack of a Ca^{2+}/calmodulin binding site suggests that *Sy* NOS protein is either regulated transcriptionally or is constitutively expressed and uses a yet unknown mechanism for activity regulation. Another notable feature of *Sy* NOS is the presence of an N-terminal globin domain with homology to bacterial flavohemoglobins. It was proposed that the globin domain acts as a NO dioxygenase that oxidizes NO to maintain NO_3^- homeostasis in the *Synechococcus* bacteria, but whether this is relevant *in vivo* remains to be determined.

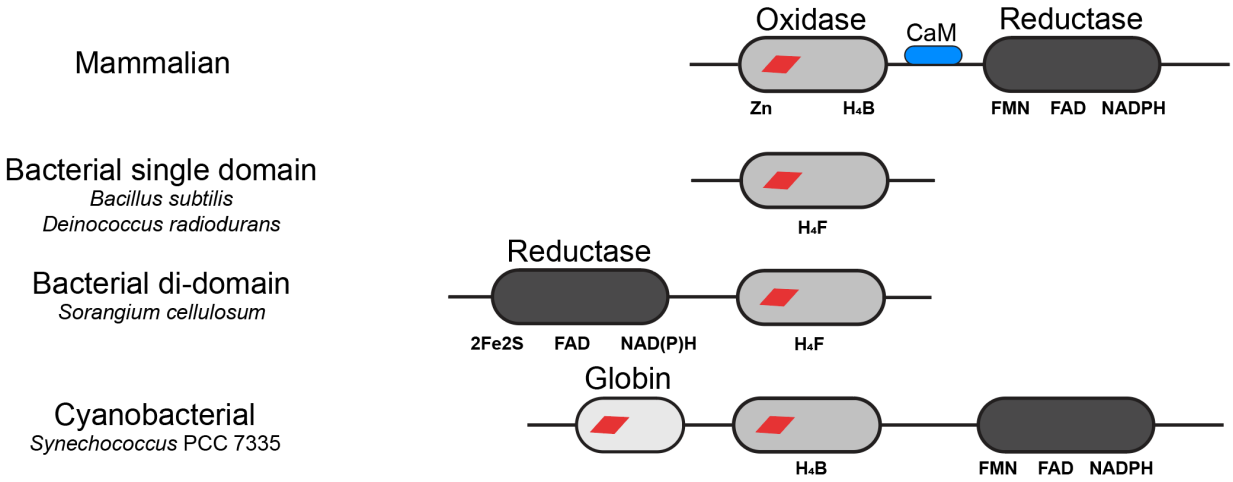


Figure 1.1. Domain architecture of nitric oxide synthases. Predicted heme binding sites are indicated with a red parallelogram. Four types of NOS are found in nature. In mammals, NOS is composed of two domains, an oxidase domain that uses a P450-like heme to catalyze the oxidation of L-arginine to produce NO, and a C-terminal reductase domain using cofactors FMN, FAD and NADPH to provide electrons to complete the catalytic cycle. Calmodulin (CaM) binding (blue) regulates the activity of mammalian NOS. In bacteria, single domain NOSs were first discovered in *Bacillus subtilis* and *Deinococcus radiodurans*, and they use a standalone reductase domain. [26,27] NOS with a reductase domain and an oxidase domain in a single polypeptide chain has also been identified in the bacteria *Sorangium cellulosum*.[28] This NOS has an iron-sulfur cluster reductase in place of FMN. Finally, a novel NOS was

identified in the cyanobacteria *Synechococcus* sp. PCC7335 that has a similar domain architecture to mammalian NOS but lacks key regulatory elements such as Zn^{2+} and Ca^{2+}/CaM binding sites. It also has a N-terminal globin of unknown function.[30]

NO sensing in eukaryotes: soluble guanylate cyclase

NO synthesized by NOS is sensed by the protein sensor of NO, sGC. sGC is an obligate dimeric protein, and each subunit of sGC is composed of four domains (Figure 1.2A). The H-NOX domain is at the N-terminus of sGC and serves as the site for gas ligand binding via the heme cofactor.[31–33] The gas binding signal is then transduced through the Per-Arnt-Sim-like domain (PAS) and the coiled coil (CC) domain. PAS-like domains are frequently found in signaling proteins and often mediate protein-protein interaction and signal transduction.[34] In the case of sGC, the PAS domain facilitates dimer formation and may be involved in interacting with molecular chaperones to aid heme insertion.[35,36] CC domains are well known for their capacity to communicate molecular signals through changes in packing and register shifting.[37] The transduction of gas binding signal through the PAS and CC domains controls the conformation of the C-terminal catalytic domain (CAT) to regulate substrate binding and catalytic activity.[38,39] The CAT domain catalyzes an intramolecular cyclization of guanosine triphosphate to yield 3',5'-cyclic guanosine monophosphate (cGMP).[40]

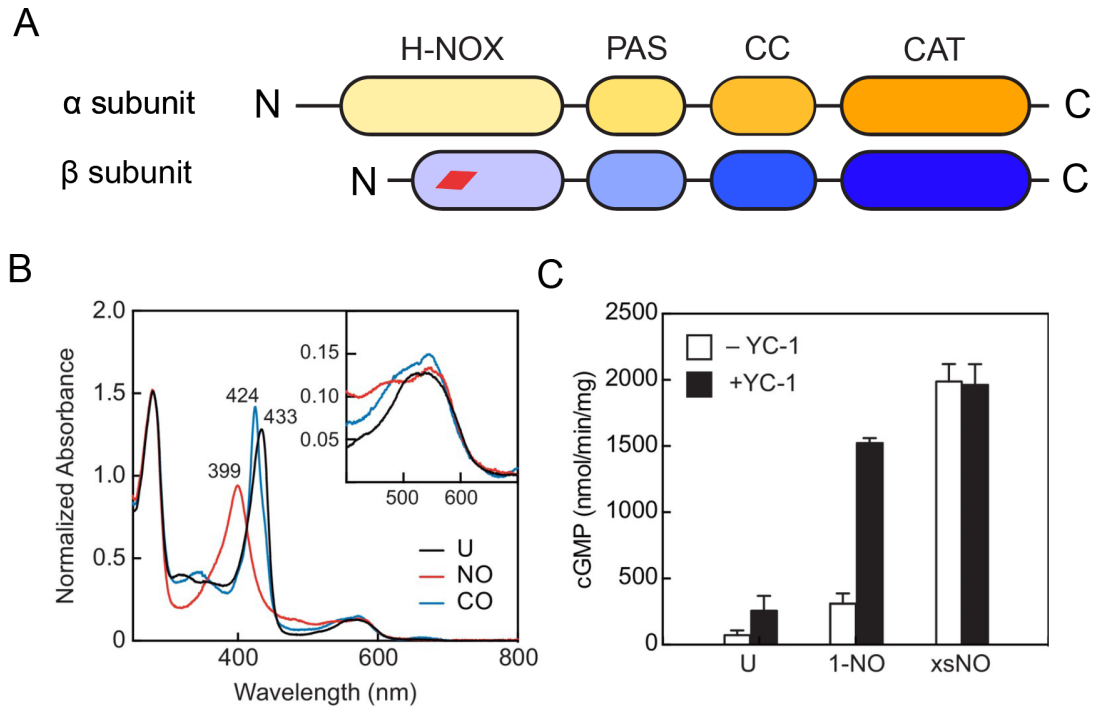


Figure 1.2. Biochemical properties of a prototypical α/β-type sGC. (A) Domain architecture of sGC. H-NOX, heme-nitric oxide-oxygen binding domain. PAS, Per-Arnt-Sim-like domain. CC, coiled coil domain. CAT, catalytic domain. (B) gas ligand binding of sGC. α/β-type sGC can bind NO and CO gas ligands.(C) Ligand-controlled activity of an α/β-type sGC. α/β-type sGCs exhibit a three-stage activation profile. Stimulators like YC-1 can activate 1-NO state sGC to xsNO-like activity. Figure adapted from ref[41].

The sGC that has been the most extensively studied is the α/β-type sGC, especially the mammalian homologs from *Homo sapiens*[39] and *Rattus norvegicus*[42] and an insect homolog from

Manduca sexta[43,44]. α/β-type sGCs, as the name suggests, are obligate heterodimers made up of α and β subunits. This endows them with a structural asymmetry that is evident throughout the length of the protein. While both the α and the β subunit harbor a predicted heme binding site by homology, only the β subunit binds heme in the dimer.[33] The reason for the lack of heme binding at the α H-NOX domain has been proposed to be the combined effect of several factors. First, heme ligation of sGC is mediated by a highly conserved histidine residue in the β H-NOX domain,[32] as well as three residues (tyrosine, serine, and arginine) arranged as the so-called YxSxR motif responsible for hydrogen bonding to the propionate side chains of heme.[45] The α subunit often lacks one or two out of the three residues which results in decreased stability of heme binding. More directly, the lack of heme binding in the α subunit could be caused by the presence of a long, disordered N-terminal loop in the α H-NOX that occupies the heme binding pocket and interferes with heme binding.[46,47] Additionally, in α/β-type sGC, the CAT domain also harbors asymmetry and contains just one, not two, catalytically competent active site.[48] The catalytic active sites of the CAT domain are located along the dimer interface, so each active site requires residues from both subunits to function.[49] Since neither the α nor the β subunits contain the full set of catalytic residues, only one of the two active sites in α/β-type sGC is capable of substrate turnover. The other site, termed the pseudosymmetric site, has been proposed to be involved with allosteric regulation by ATP.[50,51]

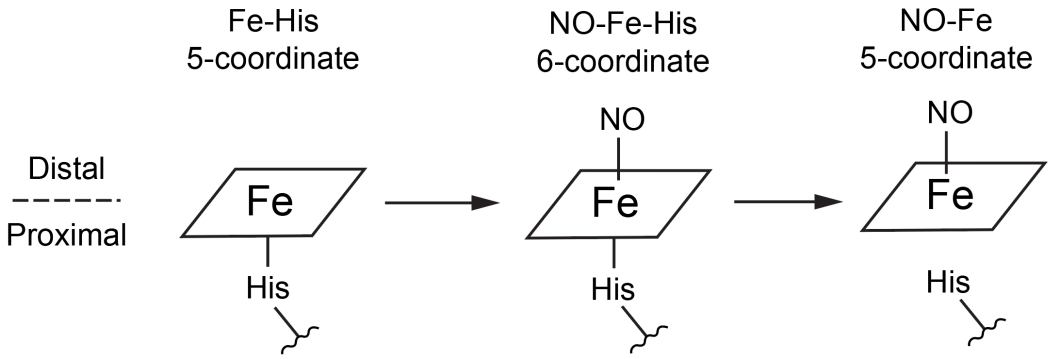


Figure 1.3. NO ligand binding to the heme cofactor of the H-NOX domain occurs through a short-lived 6-coordinate intermediate, and it involves the rupture of the proximal Fe-His bond. Figure adapted from ref[52].

When α/β-type sGC is assayed *in vitro* without any ligands bound, it exhibits a low, basal level activity (Figure 1.2B). In this state, the heme forms a 5-coordinate complex with the Fe center ligated to the histidine residue and displays a Soret band at ~430 nm in the UV-vis absorption spectrum (Figure 1.2C).[53] When NO is added to sGC, it first binds to the heme in the H-NOX domain in a 6-coordinate complex (Figure 1.3). Rupture of the Fe-histidine bond then generates a 5-coordinate complex. This complex has a Soret peak centered at 399 nm (Figure 1.2B).[53] At this stage, the sGC is moderately activated to ~15% of its full activity and is in the so-called 1-NO state (Figure 1.2C).[54,55] When NO is added in excess, the sGC can then be fully activated to the xsNO state that is greater than 25-fold over basal activity.[38] The site for interaction with excess NO is still unknown. α/β-type sGC can also bind CO in a 6-coordinate Fe(II) heme-CO complex (Figure 1.2B).[15] In this state, sGC is only moderately activated. In both the 1-NO state and the CO-bound state, α/β-type sGC can be activated to xsNO levels by small molecules known as sGC stimulators that bind to an allosteric pocket on the protein.[56] The prototypical stimulator is an indazole derivative YC-1, but fifteen stimulators and counting have been reported in the literature.[57,58] This line of research has brought forth the FDA-approved drug Riociguat that acts

through a similar mechanism as YC-1 and is used to treat pulmonary arterial hypertension.[12] The three-stage activation profile of sGC *in vitro* suggests a complex mechanism of allosteric regulation. The regulation of sGC has been studied extensively through both biochemistry and structural biology techniques and will be discussed in the following sections.

Study of H-NOX proteins

Before high-resolution structures of sGC became available, structural information was mainly derived from crystal structures of the individual domains of sGC. One especially fruitful area of research was the study of truncations of sGCs that include the N-terminal H-NOX domain as well as homolog bacterial H-NOX proteins. In bacteria, standalone H-NOX proteins were identified through homology to the H-NOX domain of eukaryotic sGCs,[59] and they can interact with cyclic-di-GMP cyclases and histidine kinases to signal quorum sensing, biofilm formation and cell growth.[60–63] Some these bacterial H-NOXs selectively bind NO under aerobic condition like sGCs[63,64] but others do not exhibit ligand selectivity and bind O_2.[65] Human pathogens, including *Vibrio cholerae*[66,67], *Legionella pneumophila*[68], and *Clostridium botulinum*,[64] are known to encode H-NOX proteins in their genomes. H-NOX signaling is associated with virulence in these pathogenic bacteria, so by itself, studying H-NOX proteins could benefit the development of new therapeutics targeting key pathways in pathogens.[69] Additionally, a great deal of our understanding of the mechanism of ligand selectivity of sGCs and how H-NOX conformational change regulates sGC activity came from structural and biochemical studies of bacterial H-NOX proteins.

The H-NOX protein is a small (<200 aa), globular protein. The secondary structure of the H-NOX protein contains seven α helices (named αA – G) and a four-stranded β sheet (named β1 – 4). The tertiary structure of H-NOX protein can be separated into two subdomains: the proximal subdomain contains the heme-ligating proximal His residue on the αF helix, αG helix, as well as the four β strands, and the distal subdomain contains αA – αE helices (Figure 1.4A). As such, the heme-binding cavity of the H-NOX protein is bisected by the heme cofactor into the distal pocket and the proximal pocket.

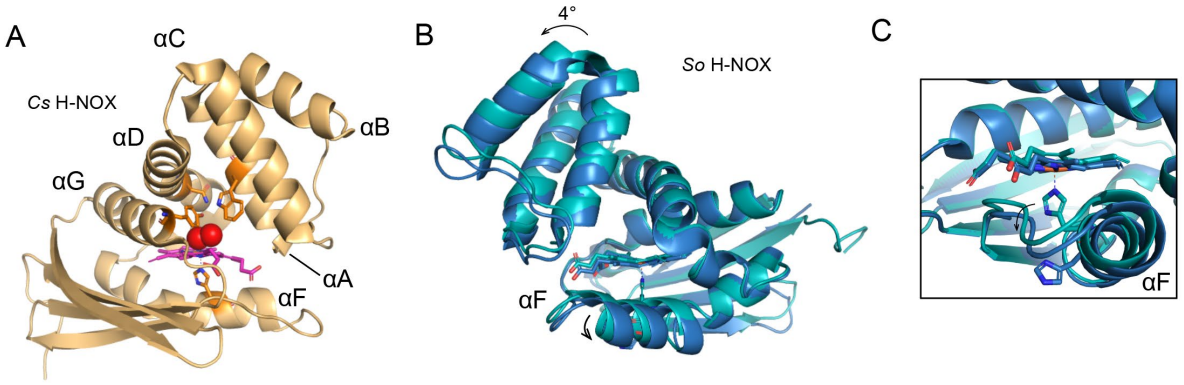


Figure 1.4. Structures of H-NOX proteins**.** (A) Distal pocket of the oxygen binding *C. subterraneous* H-NOX protein (PDB ID: 1U55). Residues within the distal pocket shown in sticks compose the hydrogen bonding network necessary for stabilizing the Fe(II) heme-O_2 complex. (B) Conformational change of the NO-specific *S. oneidensis* H-NOX protein. Teal, unliganded, PDB 4U9B; dark blue, Fe(II)-NO bound, PDB ID 4U99.[70,71] (C) Zoom in view of the conformational change taking place at the αF helix.

The ability to differentiate between NO and O_2 under aerobic conditions is crucial to the function of NO sensors, given that the concentration of O_2 is typically several orders of magnitude that of NO. In facultative anaerobic bacteria, only H-NOX proteins that are specific to NO binding were identified.[59,72] On the other hand, in obligate anaerobes, additional H-NOX proteins that can bind O_2 were identified.[59,65] The crystal structure of an O_2-binding *Caldanaerobacter subterraneous subsp. tengcongensis* H-NOX protein was solved, and a hydrogen bonding network in the distal pocket composed of residues W9, N74 and Y140 was found to stabilize the heme-O_2 complex (Figure 1.4A).[71] Y140 directly hydrogen bonds to the O_2 and was proposed to be the key residue enabling O_2 binding in *Cs* H-NOX. The Y140F variant of *Cs* H-NOX had lowered binding affinity to O_2, and a double mutant variant W9F, Y140F completely knocked out O_2 binding. Furthermore, incorporating a Tyr residue through site directed mutagenesis in a NO-specific H-NOX from *Legionella pneumophila* into the gas ligand binding pocket enabled a stable O_2 complex to form.[45] These experiments demonstrated that a distal pocket hydrogen bonding residue is necessary for O_2 binding in H-NOX domains, and if a sGC contains a tyrosine residue at the position equivalent to Y140 of *Cs* H-NOX, it can be reliably predicted to bind O_2.

As previously mentioned, in the fully active state of sGC, NO is bound to the Fe(II) heme cofactor in a 5-coordinate configuration, and the formation of this complex involves rupture of the Fe-His bond from electronic effects of NO (known as the NO *trans* effect) (Figure 1.3).[73–75] The structural detail of the Fe-His bond rupture was elucidated by the crystal structure of the NO-specific bacterial H-NOX protein from *Shewanella oneidensis* (Figure 1.4B).[70] After NO binding, structural changes include changes in heme planarity as well as a ~ 4° relative rotation between the distal and proximal subdomains hinged on the helical interface between αD and αG (Figure 1.4B). The most dramatic shift occurs in the αF helix. Residues along the αF helix rotated along the center of the helix for ~ 45°, and H103, the heme-ligating histidine, adopted a different rotamer (Figure 1.4C). In sGC, however, the conformational change of the H-NOX domain observed by Cryo-EM structures (see below) is more subtle. In *Hs* sGC, the αF helix underwent a lateral shift after the rupture of the Fe(II)-His bond, and the relative shift between the proximal and distal subdomains are less apparent.[38,39,47] Nonetheless, the structure of *So* H-NOX may represent a conformation that can be sampled by the H-NOX after NO binding.

sGC interacts with NO through non-heme interactions

The discovery of the 1-NO state of sGC has not been a straightforward process. Initially, sGC was proposed to be activated through a simple 1-step process by NO associating to the heme cofactor. The first hint that there exists a secondary sGC-NO interaction came from consecutive reports from the Marletta Lab by Stone, Zhao and Ballou[74–76] to demonstrate that the presence of excess NO affected the kinetics of heme-NO complex formation. More specifically, presence of excess NO accelerated the rate of iron-His bond rupture in the 6-coordinate to 5-coordinate Fe(II) heme-NO transition. Bellamy and colleagues[77] also made a robust case for the involvement of a second site NO interaction by coupling NO concentration to the activity of sGC. A method to achieve a constant NO concentration in solution over several minutes was developed through balancing NO release from a small molecule NO donor against a NO scavenger. Using this method, NO was titrated into a sGC cell lysate assay, and the activity of sGC under different NO concentrations was collected. The key findings of this experiment were that sGC can be fully activated under low nM concentration of NO, and that the NO-sGC concentration response curve

is best fitted with a Hill Coefficient of 2.1, suggesting a cooperative effect in NO activation of sGC. Therefore, the authors concluded that NO likely interacts with sGC with more than one site.

The involvement of a secondary NO interaction site in activating sGC was tested directly. Russwurm and Koesling[55] discovered that if sGC was first treated with excess NO, then buffer exchanged to remove the NO, its activity consistently reduced to an intermediate level higher than the NO-free, basal state. Furthermore, if NO was titrated into a sample of sGC, the formation of the NO-heme complex follows a linear trend, but the activity increase does not follow a linear trend. Cary and colleagues[50] took an alternative approach by studying sGC deactivation. The NO dissociation rate from heme and the rate of sGC inactivation were determined individually. Comparison of the two rates revealed that NO dissociation from the heme is ~160 fold lower than the deactivation rate of sGC. NO dissociation from the heme is far too slow to adequately explain the relatively rapid deactivation of sGC. Cary also demonstrated that by treating a fully active sGC sample with a NO scavenger, the sGC showing an intermediate activity state can be isolated with NO remaining bound to the heme.

The above experiments were sufficient to show that NO-sGC interaction takes place through two steps at two distinct sites, but the chemical nature of the second interaction site is still unknown. One hypothesis is that NO interacts with sGC at a non heme site, but another is that excess NO forms a transient dinitrosyl complex with the heme, and this complex subsequently converts into a proximal heme-NO complex for full activation. To address these hypotheses, Derbyshire and colleagues[78] reported that a non-physiological ligand, n-butyl isocyanide (BIC) binds to the heme strongly and was not displaced when additional NO was added. Under only BIC-bound condition, *Rn* sGC was activated ~5 fold, but with additional NO added, *Rn* sGC can be activated a further ~16 fold without NO displacing the heme-bound BIC. These results served as strong evidence that NO interacts with sGC at a non-heme site to drive sGC activation.

Besides heme, gaseous ligand interaction in proteins can also occur at a non-heme metal center or through hydrophobic binding pockets.[79,80] Yet sGC is isolated with heme as the only metal center, and CO and O_2, which are predicted to access similar binding pockets as NO, do not activate the protein. Therefore, the second interaction site is also selective for NO. Fernhoff and colleagues[54] hypothesized that a cysteine residue can interact with NO specifically, as cysteines can react with NO oxidatively to form a S-nitrosyl adduct. However, two observations led to the hypothesis that a direct, non-oxidative nucleophilic addition of a thiol to NO may be at play. First, sGC can be activated by non-heme NO without oxygen and with reductant present. Second, through gel filtration, excess NO can be removed from a sample of sGC and cause sGC to return to the ~15%-to-full activity state. If the adduct is an S-nitrosyl, it is expected to persist through gel filtration and not lead to decreased activity. The result of the thiol-NO adduct was hypothesized to be a $RSNO^{\cdot-}$, a radical anion species. Indeed, oxidatively modifying the cysteine side chains using the labeling reagent methyl methanethiosulfonate (MMTS) can inhibit the sGC to only partially stimulated, with a specific activity close to the 1-NO state, when excess NO to heme is present. The effects of MMTS were not a result of protein denaturation, irreversible modification, or disruption of NO binding at the heme cofactor. Furthermore, MMTS-treated sGC can still be activated by YC-1, implying that MMTS modification did not prevent sGC from accessing the fully active state by steric hindrance. These results presented a strong case for the involvement of cysteines in sGC activation to the fully active state.

sGC activity is regulated through conformational change

Full length sGC had been recalcitrant to protein crystallography studies, likely due to its inherent conformational flexibility. While several crystal structures of the individual domains of sGC have been solved in the past, high resolution full-length structures of sGC were only solved for the first time in 2018, thanks to the advancement of cryo-electron microscopy. Two sGC homologs were targeted in structural studies: the insect homolog of α/β-type sGC from *Manduca sexta*[38] and the mammalian homolog from *Homo sapiens*.[39,81,82] Percent sequence identity between *Ms* sGC and *Hs* sGC are 38% between the α subunits, and 61% between the β subunits. However, *Ms* sGC and *Hs* sGC show remarkably similar conformational change during activation.

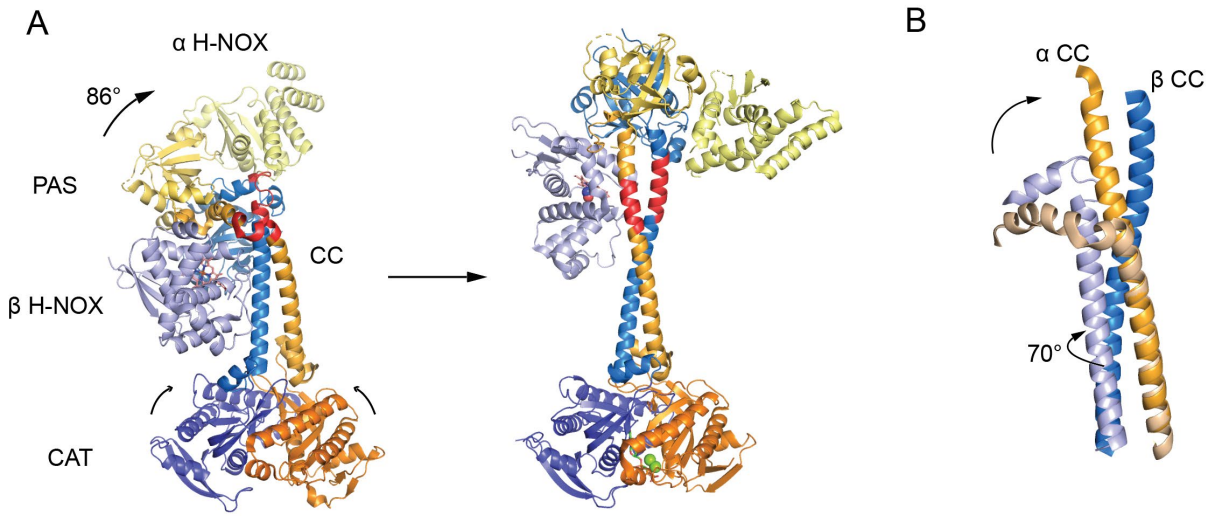


Figure 1.5. Structure of full length α/β-type sGC. (A) Structure of *H. sapiens* sGC under the inactive state and the fully active state (PDB IDs 8HBE – inactive, 8HBH – active).[82] Arrows show direction of conformational movement. The area where the coiled coil shows a bend is colored in red. (B) Structure comparison of the coiled coil in the inactive state (light brown and light blue) and the active state (orange and dark blue) by aligning the C-terminal portion of the C-terminal portion of the α subunit helices. A relative rotation of ~70° was observed.

The α and β subunits interact to form a parallel heterodimer. Under the basal activity state, α/β-type sGC displays an asymmetric conformation, where the β H-NOX domain makes direct hydrogen bonding interaction with the CC domain, while the α H-NOX domain is situated further distal (Figure 1.5A). The CC domain is bent into two halves and connected by a flexible loop (α 419-427 and β 355-359, *Ms* numbering). In the structure, the C-terminal CAT domain was solved without substrate bound. When the structure of the basal sGC is compared to the structure associated with the fully active sGC, a large-scale conformational change can be observed. Notably, the loops separating the two halves of the CC domain are straightened into α helices, resulting in an extension of the overall conformation of the protein. Furthermore, a rotation of 70° of β CC domain against α CC domain was observed (Figure 1.5B). This rotation is translated into the CAT domain and contributed to the increased separation between the two CAT domain lobes and expansion of the catalytic basket. The opening of the catalytic domain enabled more efficient substrate binding, so the Cryo-EM structure of the fully active sGC was solved with the non-hydrolysable substrate guanosine-5'-[(α,β)-methyleno]triphosphate (GMPCPP) bound. The

conformational change that takes place in the CAT domain explains the observed decrease in K_M, corresponding to a higher substrate binding affinity as well as increased catalytic efficiency.[54]

The straightening of the coiled coil is crucial for sGC activation and was studied more extensively. Kang and colleagues incorporated a proline residue into the CC domain of *Hs* sGC (α D423P or β G356P) to create a break in the coiled-coil and saw disrupted NO-sensitive activation, while an alanine incorporated in those positions did not cause a significant change.[39] Wittenborn and colleagues approached this by creating two variants, one where the loop area was replaced by a disordered linker composed of primarily glycine and serine residues (GS linker) to impart greater structural flexibility, and another that a leucine zipper motif was installed to render the CC domain constitutively extended.[83] As expected, both variants lost the ability to respond to NO. The GS linker variant displays basal activity even when NO is present in molar excess, and the leucine zipper variant is constitutively active. Both Kang's and Wittenborn's work illustrate the importance of the finely tuned conformational flexibility of the CC domain in regulating intramolecular signal transduction in sGC.

Another useful technique applied to studying the conformational change of sGC has been Small Angle X-ray Scattering (SAXS). Compared to high resolution structural techniques such as X-ray crystallography and Cryo-EM, SAXS is a relatively low-resolution technique, its resolution limited to ~10 Å.[84] However, as SAXS experiments measure scattering of X-ray by biomolecules in solution, it does not require specialized sample preparation steps and is highly adaptable to studying large, complex biomolecules such as sGC under native-like environments.[85] Protein SAXS experiments can be carried out in a static X-ray cell or flow cell, or coupled to a chromatography setup, often size exclusion chromatography (SEC). The benefit to using chromatography-coupled SAXS is that a heterogeneous sample can be separated into individual components to facilitate analysis. Using SEC-SAXS, Horst and Yokom showed that the 1-NO state of *Ms* sGC is best modeled as an ensemble of two distinct conformational states.[38] Under NO-free conditions, *Ms* sGC can be modeled exclusively as the unliganded, compact conformation. In comparison, the 1-NO state was best modeled with 72% protein in the compact conformation and the rest 28% in a partially extended conformation. This suggests that NO binding at the heme cofactor shifts the conformational equilibrium towards greater conformational extension and could help explain why *Ms* sGC shows ~15% full activity under the 1-NO state.

Homodimeric sGC: beyond NO sensing

While α/β-type sGCs are the most relevant to human physiology, and thus have received the most attention, another class of sGC harbors great biochemical diversity but has not been as extensively studied. Homodimeric sGCs are found in various groups of organisms including many animals;[86] they are also found in non-animal eukaryotes, with notable examples such as the algal sGCs from *Chlamydomonas reinhardtii*[41,87] and more recently, sGCs from choanoflagellates.[88] Homodimeric sGCs are similarly obligate dimers for activity and share the domain architecture of α/β-type sGCs. However, each monomer of a homodimeric sGC contains all necessary residues for catalysis in the CAT domain, so in theory, and frequently in practice, they can be active as homodimers.[89]

Homodimeric sGC can exhibit distinct biochemical properties that differ from α/β-type sGC. For example, Horst and colleagues[41] reported that the sGC Cyg-11 from the algae *C. reinhardtii* is activated by CO binding to ~4 fold its basal activity, compared to a ~2 fold activation when NO is bound. Furthermore, *Cr* Cyg-11 exhibits distinct gas ligand binding kinetics. NO dissociation of *Cr* Cyg-11 is 3 orders of magnitude faster compared to other characterized H-NOX domains, and the affinity for NO as measured by the dissociation constant K_D is 2 orders of magnitude lower than typical α/β-type sGCs such as rat sGC. In a separate *in vivo* study, CO was proposed to be generated from heme oxidation by heme oxygenase I, and CO can upregulate several of the genes involved in iron acquisition and recycling.[90] Given this, Horst and colleagues proposed that *Cr* Cyg-11 is a physiological sensor for CO that mediates iron-deficiency response in its native organism.

Homodimeric sGCs also include those that can bind O_2. The biological roles of predicted O_2-binding sGC in animals have been explored *in vivo*. In *Caenorhabditis elegans*, sGCs were shown to express by neurons and mediate aerotaxis.[91–93] The N-terminal portion of the sGC Gcy-35 that contains the H-NOX domain was expressed and shown to bind O_2. However, Gcy-35 could not be expressed in full-length in a heterologous host, which precluded directly associating the ligand-induced activity of Gcy-35 to the phenotype.[91] In *Drosophila melanogaster*, three predicted O_2-binding sGCs were identified and named Gyc-88E, Gyc-89Da and Gyc-89Db.[94,95] In vitro assays demonstrated that Gyc-89Da and Gyc-89Db cannot form active homodimers.[96] While Gyc-88E is active as a homodimer, it is frequently co-expressed with Gyc-89Da and Gyc-89Db, suggesting that heterodimers are likely the physiologically relevant form of these proteins.[96,97] Morton reported quantification of cGMP production *in cellulo* through transient expression of genes encoding the three sGCs and discovered that cGMP production increased under anaerobic conditions.[95]

Huang and colleagues reported *in vitro* characterization of *Dm* Gyc-88E.[98] Gyc-88E remains the only O_2-binding sGC that has been reported in the literature to be characterized *in vitro* in purified form. Major differences between Gyc-88E and α/β-type sGC include gas ligand binding and ligand-induced activity change. Like Gcy-35 from *C. elegans*, Gyc-88E binds O_2 to form a 6-coordinate Fe(II)-O_2 complex; and noticeably, it binds NO to form a 6-coordinate Fe(II)-NO complex as well, a marked difference from the 5-coordinate Fe(II)-NO complex formed by α/β-type sGC. Additionally, Gyc-88E is inhibited by gas ligand binding (including O_2), which matches the results observed by Morton in *in vivo* studies.[95]

Expanding the scope of sGC-mediated NO signaling

Beyond vertebrates and insects, NO signaling has been found to control behaviors in a wide range of animals including bilaterians that contain ascidians (sea squirts), mollusks and echinoderms (sea star, sea urchin), to name a few, and non-bilaterians to include sponges, cnidaria (jellyfish, sea anemone), and placozoans.[99–106] In these organisms, NO signaling controls a wide range of behaviors that include motor function and feeding, collective contraction, flagellar beating and larval metamorphosis. Granted, NO signaling can occur through a few different pathways, including protein S-nitrosation and interaction with hemoproteins that are not sGC.[107] Therefore, two pathways with strong evidence for the involvement of sGC are highlighted here.

In the jellyfish *Aglantha digitale*, NO-cGMP signaling controls swimming behavior that is important for its feeding cycle.[102] The authors first established neuronal localization of NOS by NADPH-diaphorase staining and tested NO production through inhibition of NOS and detection of NO oxidation products nitrite and nitrate. Both exogenous NO and the cell-permeable cGMP analog, 8-Br-cGMP can cause an increase in swimming frequency, supporting a role for the involvement of a NO-activated sGC. The genome of *A. digitale* became available in 2023, and through a BLAST search, sGCs that are predicted to be NO specific are identified.[108]

NO-cGMP signaling is also used by sponges to control coordinated contraction of the aquiferous canal system for fluid exchange. In the freshwater sponge *Ephydatia muelleri*, the activity of NOS was supported by NADPH-diaphorase staining.[100] Furthermore, NO-triggered accumulation of cGMP can be observed in contractile cells through immunofluorescent staining targeting cGMP. NO-controlled behaviors have also been identified in a few other sponge species including *Tethya wilhelma*,[109] and *Amphimedon queenslandica*,[110] and in *Spongilla lacustris* coexpression of NOS and sGC has been observed through RNA-seq experiments.[111] These results highlight the potential for new discoveries with a broadened scope of NO-cGMP signaling.

Given that NO-cGMP signaling be found in some of the earliest diverging lineages of animals, cnidarians and sponges, an interesting question now arises: how did NO signaling evolve in early animals? To answer this, we turned our attention to choanoflagellates. Choanoflagellates are the closest unicellular relatives of animals. Multicellularity evolved independently for animals, plants and fungi, so did molecular pathways involved in regulating intercellular signaling, adhesion, motility, and cell type differentiation that are crucial in multicellularity.[112] Several species of choanoflagellates have been developed as model organisms, so the molecular pathways that are conserved across choanoflagellates and animals can be studied in detail.[113] These studies will help reconstruct the role of these pathways in the last unicellular ancestor of animals.[114] Of special interest to the work presented in this thesis is the colonial choanoflagellate *Choanoeca flexa*. *C. flexa* is an emerging model organism for studying multicellular behavior in choanoflagellates. Discovered in 2018 by members of the King Lab, Brunet, Linden and Larson in the tide pools on the shores of Curaçao, *C. flexa* was named after an interesting phenotype it displays.[115] *C. flexa* forms flexible sheet colonies that can switch between two distinct morphologies. When the flagella of the colonies face outward, motility is favored. While the flagella face inward, feeding is facilitated. By regulating the feeding of nutrients, the authors discovered that a rhodopsin-linked cGMP specific phosphodiesterase is involved in regulating cellular cGMP levels that then regulates the transition between the two colony morphologies. This discovery raised hypotheses of *C. flexa* using sGCs for generating the cGMP in this pathway.

Thesis research

The focus of this dissertation is on the study of sGCs from the choanoflagellate species *Choanoeca flexa*. Current knowledge of sGC has mostly originated from studies of α/β-type sGCs from mammals and insects. Progress in this area has been weighed down by the low heterologous expression yield of these sGCs and the complicated cloning process associated with insect cell expression. Furthermore, with newfound knowledge of the conformational change in α/β-type sGCs, the question of whether activity regulation through conformational extension is a common theme across diverse sGC homologs is becoming increasingly relevant.

Chapter 2 describes work carried out in collaboration with Dr. Josean Reyes-Rivera and Dr. Thibaut Brunet from the lab of Prof. Nicole King. Experiments were carried out *in vivo* to characterize a NO sensitive phenotype in *C. flexa*. *C. flexa* forms multicellular sheet colonies that undergo collective contraction, and this process is regulated through cGMP signaling. A BLAST search revealed that *C. flexa* contains both NOS and several sGCs in its genome, suggesting the presence of a NO-sGC-cGMP signaling axis. Coincubation of a NO donor with a colonial *C. flexa* culture demonstrated that NO triggers collective contraction in *C. flexa*. Through sequence alignment, *Cf* sGC1 was identified as a NO-specific sGC and was predicted to play a key role in regulating NO-mediated colony contraction. Gas ligand binding properties of *Cf* sGC1 were reported in the preliminary characterization in this chapter. *Cf* sGC1 is also activated by NO binding, implying that the NO-mediated colony contraction is partly regulated by *Cf* sGC1. When sGCs in *C. flexa* were inhibited by the pan-sGC inhibitor ODQ, colony contraction was still initiated by NO addition, but the duration greatly decreased. Therefore, NO-mediated colony contraction in *C. flexa* likely involves more signaling partners beyond sGCs.

Chapter 3 reports a detailed biochemical characterization of *Cf* sGC1. *Cf* sGC1 is the first homodimeric sGC reported that exhibits a 3-stage activation profile of α/β-type sGC. This raised questions about the mechanism of activity regulation in *Cf* sGC1. Ligand-bound activity of *Cf* sGC1 was reanalyzed to include CO. CO activates *Cf* sGC1 but only to 1-NO levels of activity. Despite having two possible heme binding sites, *Cf* sGC1 only binds one heme per homodimer. Small angle X-ray scattering (SAXS) was used to study the oligomeric state and the conformation of *Cf* sGC1. *Cf* sGC1 is an asymmetric homodimer, and during activation, a conformational extension like that of α/β-type sGC was observed.

Chapter 4 explores the mechanism of non-heme NO activation in sGCs using *Cf* sGC1 as a model system. Both α/β-type sGC and *Cf* sGC1 require excess NO to reach full activity, suggesting the involvement of a second NO interaction site. Cysteines have been proposed to be the site of non-heme NO interaction during α/β-type sGC activation, and cysteine variants of *Cf* sGC1 at conserved positions were prepared to probe the roles of cysteines in regulating *Cf* sGC1 activity.

Chapter 5 discusses the characterization of an O_2 binding sGC from *C. flexa*. *Cf* sGC4 reported in this chapter is only the second of its kind to be characterized *in vitro*. *Cf* sGC4 binds NO, CO and O_2. Ligand binding leads to the formation of a 6-coordinate Fe(II)-heme complex. Although *Cf* sGC4 is inhibited by gas ligand binding in activity assays, a conformational change was not observed through SAXS. This raised the possibility that *Cf* sGC4 is regulated by a different mechanism from that observed for α/β-type sGC and *Cf* sGC1 in chapter 3.

References

(1) Martínez-Ruiz, A.; Cadenas, S.; Lamas, S. Nitric Oxide Signaling: Classical, Less Classical, and Nonclassical Mechanisms. *Free. Radic. Biol. Med.* **2011**, *51* (1), 17–29

(2) Palmer, R. M. J.; Ferrige, A. G.; Moncada, S. Nitric Oxide Release Accounts for the Biological Activity of Endothelium-Derived Relaxing Factor. *Nature* **1987**, *327* (6122), 524–526

(3) Ignarro, L. J.; Byrns, R. E.; Buga, G. M.; Wood, K. S. Endothelium-Derived Relaxing Factor from Pulmonary Artery and Vein Possesses Pharmacologic and Chemical Properties Identical to Those of Nitric Oxide Radical. *Circ. Res.* **1987**, *61* (6), 866–879

(4) Arnold, W. P.; Mittal, C. K.; Katsuki, S.; Murad, F. Nitric Oxide Activates Guanylate Cyclase and Increases Guanosine 3',5'-Cyclic Monophosphate Levels in Various Tissue Preparations. *Proc. Natl. Acad. Sci. U S A* **1977**, *74* (8), 3203–3207

(5) Marletta, M. A.; Yoon, P. S.; Iyengar, R.; Leaf, C. D.; Wishnok, J. S. Macrophage Oxidation of L-Arginine to Nitrite and Nitrate: Nitric Oxide Is an Intermediate. *Biochemistry* **1988**, *27* (24), 8706–8711

(6) Arancio, O.; Kiebler, M.; Lee, C. J.; Lev-Ram, V.; Tsien, R. Y.; Kandel, E. R.; Hawkins, R. D. Nitric Oxide Acts Directly in the Presynaptic Neuron to Produce Long-Term Potentiation in Cultured Hippocampal Neurons. *Cell* **1996**, *87* (6), 1025–1035

(7) Aso, Y.; Ray, R. P.; Long, X.; Bushey, D.; Cichewicz, K.; Ngo, T. T.; Sharp, B.; Christoforou, C.; Hu, A.; Lemire, A.; Tillberg, P.; Hirsh, J.; Litwin-Kumar, A.; Rubin, G. M. Nitric Oxide Acts as a Cotransmitter in a Subset of Dopaminergic Neurons to Diversify Memory Dynamics. *eLife* **2019**, 9:e64094

(8) Alexander, B. The Role of Nitric Oxide in Hepatic Metabolism. *Nutrition* **1998**, *14* (4), 376–390

(9) Sharma, J. N.; Al-Omran, A.; Parvathy, S. S. Role of Nitric Oxide in Inflammatory Diseases. *Inflammopharmacology* **2007**, *15* (6), 252–259

(10) Steinert, J. R.; Chernova, T.; Forsythe, I. D. Nitric Oxide Signaling in Brain Function, Dysfunction, and Dementia. *The Neuroscientist* **2010**, *16* (4), 435–452

(11) Stasch, J. P.; Pacher, P.; Evgenov, O. V. Soluble Guanylate Cyclase as an Emerging Therapeutic Target in Cardiopulmonary Disease. *Circulation* **2011**, *123* (20), 2263–2273

(12) Conole, D.; Scott, L. J. Riociguat: First Global Approval. *Drugs* **2013**, *73* (17), 1967–1975

(13) Knowles, R. G.; Moncada, S. Nitric Oxide Synthases in Mammals. *Biochemical Journal* **1994**, *298* (2), 249–258

(14) Knowles, R. G.; Palacios, M.; Palmer, R. M. J.; Moncada, S. Formation of Nitric Oxide from L-Arginine in the Central Nervous System: A Transduction Mechanism for Stimulation of the Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U S A* **1989**, *86* (13), 5159–5162

(15) Stone, J. R.; Marletta, M. A. Soluble Guanylate Cyclase from Bovine Lung: Activation with Nitric Oxide and Carbon Monoxide and Spectral Characterization of the Ferrous and Ferric States. *Biochemistry* **1994**, *33* (18), 5636–5640

(16) Tsai, E. J.; Kass, D. A. Cyclic GMP Signaling in Cardiovascular Pathophysiology and Therapeutics. *Pharmacol. Ther.* **2009**, *122* (3), 216–238

(17) Forstermann, U.; Sessa, W. C. Nitric Oxide Synthases: Regulation and Function. *Eur. Heart J.* **2012**, *33* (7), 829–837

(18) Campbell, M. G.; Smith, B. C.; Potter, C. S.; Carragher, B.; Marletta, M. A. Molecular Architecture of Mammalian Nitric Oxide Synthases. *Proc. Natl. Acad. Sci. U S A* **2014**, *111* (35), 3614–3623

(19) Hurshman, A. R.; Marletta, M. A. Reactions Catalyzed by the Heme Domain of Inducible Nitric Oxide Synthase: Evidence for the Involvement of Tetrahydrobiopterin in Electron Transfer. *Biochemistry* **2002**, *41* (10), 3439–3456

(20) Hemmens, B.; Goessler, W.; Schmidt, K.; Mayer, B. Role of Bound Zinc in Dimer Stabilization but Not Enzyme Activity of Neuronal Nitric-Oxide Synthase. *J. Biol. Chem* **2000**, *275* (46), 35786–35791

(21) Mitchell, D. A.; Erwin, P. A.; Michel, T.; Marletta, M. A. S-Nitrosation and Regulation of Inducible Nitric Oxide Synthase. *Biochemistry* **2005**, *44* (12), 4636–4647

(22) Lamas, S.; Marsden, P. A.; Li, G. K.; Tempst, P.; Michel, T. Endothelial Nitric Oxide Synthase: Molecular Cloning and Characterization of a Distinct Constitutive Enzyme Isoform. *Proc. Natl. Acad. Sci. U S A* **1992**, *89* (14), 6348–6352

(23) Bredt, D. S.; Snyder, S. H. Nitric Oxide, a Novel Neuronal Messenger. *Neuron* **1992**, *8* (1), 3–11

(24) Cho, H. J.; Xie, Q. W.; Calaycay, J.; Mumford, R. A.; Swiderek, K. M.; Lee, T. D.; Nathan, C. Calmodulin Is a Subunit of Nitric Oxide Synthase from Macrophages. *J. Exp. Med*. **1992**, *176* (2), 599–604

(25) Gribovskaja, I.; Brownlow, K. C.; Dennis, S. J.; Rosko, A. J.; Marletta, M. A.; Stevens-Truss, R. Calcium-Binding Sites of Calmodulin and Electron Transfer by Inducible Nitric Oxide Synthase. *Biochemistry* **2005**, *44* (20), 7593–7601

(26) Gusarov, I.; Starodubtseva, M.; Wang, Z.-Q.; McQuade, L.; Lippard, S. J.; Stuehr, D. J.; Nudler, E. Bacterial Nitric-Oxide Synthases Operate without a Dedicated Redox Partner. *J. Biol. Chem.* **2008**, *283* (19), 13140–13147

(27) Adak, S.; Bilwes, A. M.; Panda, K.; Hosfield, D.; Aulak, K. S.; McDonald, J. F.; Tainer, J. A.; Getzoff, E. D.; Crane, B. R.; Stuehr, D. J. Cloning, Expression, and Characterization of a Nitric Oxide Synthase Protein from *Deinococcus radiodurans*. *Proc. Natl. Acad. Sci. U S A* **2002**, *99* (1), 107–112

(28) Agapie, T.; Suseno, S.; Woodward, J. J.; Stoll, S.; Britt, R. D.; Marletta, M. A. NO Formation by a Catalytically Self-Sufficient Bacterial Nitric Oxide Synthase from *Sorangium cellulosum*. *Proc. Natl. Acad. Sci. U S A* **2009**, *106* (38), 16221–16226

(29) Reece, S. Y.; Woodward, J. J.; Marletta, M. A. Synthesis of Nitric Oxide by the NOS-like Protein from *Deinococcus radiodurans* : A Direct Role for Tetrahydrofolate. *Biochemistry* **2009**, *48* (23), 5483–5491

(30) Picciano, A. L.; Crane, B. R. A Nitric Oxide Synthase–like Protein from *Synechococcus* Produces NO/NO_3^- from l-Arginine and NAPDH in a Tetrahydrobiopterin- and Ca^{2+}-Dependent Manner. *J. Biol. Chem.* **2019**, *294* (27), 10708–10719

(31) Wedel, B.; Humbert, P.; Harteneck, C.; Foerster, J.; Malkewitz, J.; Böhme, E.; Schultz, G.; Koesling, D. Mutation of His-105 in the Beta 1 Subunit Yields a Nitric Oxide-Insensitive Form of Soluble Guanylyl Cyclase. *Proc. Natl. Acad. Sci. U S A* **1994**, *91* (7), 2592–2596

(32) Zhao, Y.; Schelvis, J. P. M.; Babcock, G. T.; Marletta, M. A. Identification of Histidine 105 in the B1 Subunit of Soluble Guanylate Cyclase as the Heme Proximal Ligand. *Biochemistry* **1998**, *37* (13), 4502–4509

(33) Zhao, Y.; Marletta, M. A. Localization of the Heme Binding Region in Soluble Guanylate Cyclase. *Biochemistry* **1997**, *36* (50), 15959–15964

(34) Montfort, W. R. Per-ARNT-Sim Domains in Nitric Oxide Signaling by Soluble Guanylyl Cyclase. *J. Mol. Biol.* **2024**, *436* (3), 168235

(35) Sarkar, A.; Dai, Y.; Haque, M. M.; Seeger, F.; Ghosh, A.; Garcin, E. D.; Montfort, W. R.; Hazen, S. L.; Misra, S.; Stuehr, D. J. Heat Shock Protein 90 Associates with the Per-Arnt-Sim Domain of Heme-Free Soluble Guanylate Cyclase. *J. Biol. Chem.* **2015**, *290* (35), 21615–21628

(36) Purohit, R.; Weichsel, A.; Montfort, W. R. Crystal Structure of the Alpha Subunit PAS Domain from Soluble Guanylyl Cyclase. *Protein Science* **2013**, *22* (10), 1439–1444

(37) Truebestein, L.; Leonard, T. A. Coiled-coils: The Long and Short of It. *BioEssays* **2016**, *38* (9), 903–916

(38) Horst, B. G.; Yokom, A. L.; Rosenberg, D. J.; Morris, K. L.; Hammel, M.; Hurley, J. H.; Marletta, M. A. Allosteric Activation of the Nitric Oxide Receptor Soluble Guanylate Cyclase Mapped by Cryo-Electron Microscopy. *eLife* **2019**, 8:e50634

(39) Kang, Y.; Liu, R.; Wu, J.-X.; Chen, L. Structural Insights into the Mechanism of Human Soluble Guanylate Cyclase. *Nature* **2019**, *574* (7777), 206–210

(40) Winger, J. A.; Marletta, M. A. Expression and Characterization of the Catalytic Domains of Soluble Guanylate Cyclase: Interaction with the Heme Domain. *Biochemistry* **2005**, *44* (10), 4083–4090

(41) Horst, B. G.; Stewart, E. M.; Nazarian, A. A.; Marletta, M. A. Characterization of a Carbon Monoxide-Activated Soluble Guanylate Cyclase from *Chlamydomonas reinhardtii*. *Biochemistry* **2019**, *58* (17), 2250–2259

(42) Braughler, J. M.; Mittal, C. K.; Murad, F. Purification of Soluble Guanylate Cyclase from Rat Liver. *Proc. Natl. Acad. Sci. U S A* **1979**, *76* (1), 219–222

(43) Hu, X.; Murata, L. B.; Weichsel, A.; Brailey, J. L.; Roberts, S. A.; Nighorn, A.; Montfort, W. R. Allostery in Recombinant Soluble Guanylyl Cyclase from *Manduca sexta*. *J. Biol. Chem.* **2008**, *283* (30), 20968–20977

(44) Gibson, N. J.; Nighorn, A. Expression of Nitric Oxide Synthase and Soluble Guanylyl Cyclase in the Developing Olfactory System of *Manduca sexta*. *J. Comp. Neurol.* **2000**, *422* (2), 191–205

(45) Boon, E. M.; Huang, S. H.; Marletta, M. A. A Molecular Basis for NO Selectivity in Soluble Guanylate Cyclase. *Nat.Chem. Biol.* **2005**, *1* (1), 53–59

(46) Zhong, F.; Pan, J.; Liu, X.; Wang, H.; Ying, T.; Su, J.; Huang, Z.-X.; Tan, X. A Novel Insight into the Heme and NO/CO Binding Mechanism of the Alpha Subunit of Human Soluble Guanylate Cyclase. *JBIC Journal of Biological Inorganic Chemistry* **2011**, *16* (8), 1227–1239

(47) Wittenborn, E. C.; Marletta, M. A. Structural Perspectives on the Mechanism of Soluble Guanylate Cyclase Activation. *Int. J. Mol. Sci*. **2021**, *22* (11), 5439

(48) Childers, K. C.; Garcin, E. D. Structure/Function of the Soluble Guanylyl Cyclase Catalytic Domain. *Nitric Oxide* **2018**, *77*, 53–64

(49) Allerston, C. K.; von Delft, F.; Gileadi, O. Crystal Structures of the Catalytic Domain of Human Soluble Guanylate Cyclase. *PLoS One* **2013**, *8* (3), e57644

(50) Cary, S. P. L.; Winger, J. A.; Marletta, M. A. Tonic and Acute Nitric Oxide Signaling through Soluble Guanylate Cyclase Is Mediated by Nonheme Nitric Oxide, ATP, and GTP. *Proc. Natl. Acad. Sci. U S A* **2005**, *102* (37), 13064–13069

(51) Derbyshire, E. R.; Fernhoff, N. B.; Deng, S.; Marletta, M. A. Nucleotide Regulation of Soluble Guanylate Cyclase Substrate Specificity. *Biochemistry* **2009**, *48* (31), 7519–7524

(52) Guo, Y.; Suess, D. L. M.; Herzik, M. A.; Iavarone, A. T.; Britt, R. D.; Marletta, M. A. Regulation of Nitric Oxide Signaling by Formation of a Distal Receptor–Ligand Complex. *Nat. Chem. Biol.* **2017**, *13* (12), 1216–1221

(53) Gerzer, R.; Böhme, E.; Hofmann, F.; Schultz, G. Soluble Guanylate Cyclase Purified from Bovine Lung Contains Heme and Copper. *FEBS Lett* **1981**, *132* (1), 71–74

(54) Fernhoff, N. B.; Derbyshire, E. R.; Marletta, M. A. A Nitric Oxide/Cysteine Interaction Mediates the Activation of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U S A* **2009**, *106* (51), 21602–21607

(55) Russwurm, M.; Koesling, D. NO Activation of Guanylyl Cyclase. *EMBO J.* **2004**, *23* (22), 4443–4450

(56) Stones, J. R.; Marletta, M. A. Synergistic Activation of Soluble Guanylate Cyclase by YC-1 and Carbon Monoxide: Implications for the Role of Cleavage of the Iron-Histidine Bond during Activation by Nitric Oxide. *Chem. Biol.* **1998**, *5* (5), 255–261

(57) Ko, F.; Wu, C.; Kuo, S.; Lee, F.; Teng, C. YC-1, a Novel Activator of Platelet Guanylate Cyclase. *Blood* **1994**, *84* (12), 4226–4233

(58) Buys, E. S.; Zimmer, D. P.; Chickering, J.; Graul, R.; Chien, Y. T.; Profy, A.; Hadcock, J. R.; Masferrer, J. L.; Milne, G. T. Discovery and Development of next Generation sGC Stimulators with Diverse Multidimensional Pharmacology and Broad Therapeutic Potential. *Nitric Oxide* **2018**, *78*, 72–80

(59) Iyer, L. M.; Anantharaman, V.; Aravind, L. Ancient Conserved Domains Shared by Animal Soluble Guanylyl Cyclases and Bacterial Signaling Proteins. *BMC Genomics* **2003**, *4* (1), 5

(60) Hengge, R. Principles of C-Di-GMP Signalling in Bacteria. *Nat. Rev. Microbiol.* **2009**, *7* (4), 263–273

(61) Skerker, J. M.; Prasol, M. S.; Perchuk, B. S.; Biondi, E. G.; Laub, M. T. Two-Component Signal Transduction Pathways Regulating Growth and Cell Cycle Progression in a Bacterium: A System-Level Analysis. *PLoS Biol.* **2005**, *3* (10), e334

(62) Liu, N.; Xu, Y.; Hossain, S.; Huang, N.; Coursolle, D.; Gralnick, J. A.; Boon, E. M. Nitric Oxide Regulation of Cyclic di-GMP Synthesis and Hydrolysis in *Shewanella Woodyi*. *Biochemistry* **2012**, *51* (10), 2087–2099

(63) Price, M. S.; Chao, L. Y.; Marletta, M. A. *Shewanella Oneidensis* MR-1 H-NOX Regulation of a Histidine Kinase by Nitric Oxide. *Biochemistry* **2007**, *46* (48), 13677–13683

(64) Nioche, P.; Berka, V.; Vipond, J.; Minton, N.; Tsai, A.-L.; Raman, C. S. Femtomolar Sensitivity of a NO Sensor from *Clostridium botulinum*. *Science* **2004**, *306* (5701), 1550–1553

(65) Karow, D. S.; Pan, D.; Tran, R.; Pellicena, P.; Presley, A.; Mathies, R. A.; Marletta, M. A. Spectroscopic Characterization of the Soluble Guanylate Cyclase-like Heme Domains from *Vibrio cholerae* and *Thermoanaerobacter tengcongensis*. *Biochemistry* **2004**, *43* (31), 10203–10211

(66) Waters, C. M.; Lu, W.; Rabinowitz, J. D.; Bassler, B. L. Quorum Sensing Controls Biofilm Formation in *Vibrio cholerae* through Modulation of Cyclic di-GMP Levels and Repression of *VpsT*. *J. Bacteriol.* **2008**, *190* (7), 2527–2536

(67) Hammer, B. K.; Bassler, B. L. Distinct Sensory Pathways in *Vibrio cholerae* El Tor and Classical Biotypes Modulate Cyclic Dimeric GMP Levels To Control Biofilm Formation. *J. Bacteriol.* **2009**, *191* (1), 169–177

(68) Carlson, H. K.; Vance, R. E.; Marletta, M. A. H-NOX Regulation of c-di-GMP Metabolism and Biofilm Formation in *Legionella pneumophila*. *Mol. Microbiol.* **2010**, *77* (4), 930–942

(69) Lee-Lopez, C.; Yukl, E. H-NOX Proteins in the Virulence of Pathogenic Bacteria. *Biosci. Rep.* **2022**, *42* (1)

(70) Rao, M.; Herzik, M. A.; Iavarone, A. T.; Marletta, M. A. Nitric Oxide-Induced Conformational Changes Govern H-NOX and Histidine Kinase Interaction and Regulation in *Shewanella oneidensis*. *Biochemistry* **2017**, *56* (9), 1274–1284

(71) Pellicena, P.; Karow, D. S.; Boon, E. M.; Marletta, M. A.; Kuriyan, J. Crystal Structure of an Oxygen-Binding Heme Domain Related to Soluble Guanylate Cyclases. *Proc. Natl. Acad. Sci. U S A* **2004**, *101* (35), 12854–12859

(72) Boon, E. M.; Davis, J. H.; Tran, R.; Karow, D. S.; Huang, S. H.; Pan, D.; Miazgowicz, M. M.; Mathies, R. A.; Marletta, M. A. Nitric Oxide Binding to Prokaryotic Homologs of the Soluble Guanylate Cyclase B1 H-NOX Domain. *J. Biol. Chem.* **2006**, *281* (31), 21892–21902

(73) Hunt, A. P.; Lehnert, N. Heme-Nitrosyls: Electronic Structure Implications for Function in Biology. *Acc. Chem. Res.* **2015**, *48* (7), 2117–2125

(74) Ballou, D. P.; Zhao, Y.; Brandish, P. E.; Marletta, M. A. Revisiting the Kinetics of Nitric Oxide (NO) Binding to Soluble Guanylate Cyclase: The Simple NO-Binding Model Is Incorrect. *Proc. Natl. Acad. Sci. U S A* **2002**, *99* (19), 12097–12101

(75) Zhao, Y.; Brandish, P. E.; Ballou, D. P.; Marletta, M. A. A Molecular Basis for Nitric Oxide Sensing by Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U S A* **1999**, *96* (26), 14753–14758

(76) Stone, J. R.; Marletta, M. A. Spectral and Kinetic Studies on the Activation of Soluble Guanylate Cyclase by Nitric Oxide. *Biochemistry* **1996**, *35* (4), 1093–1099

(77) Bellamy, T. C.; Garthwaite, J. The Receptor-like Properties of Nitric Oxide-Activated Soluble Guanylyl Cyclase in Intact Cells. In *Guanylate Cyclase*; Springer US: Boston, MA, 2002; pp 165–176

(78) Derbyshire, E. R.; Marletta, M. A. Butyl Isocyanide as a Probe of the Activation Mechanism of Soluble Guanylate Cyclase: Investigating the Role of Non-Heme Nitric Oxide. *J. Biol. Chem.* **2007**, *282* (49), 35741–35748

(79) Aono, S. Metal-Containing Sensor Proteins Sensing Diatomic Gas Molecules. *Dalton Trans.* **2008**, No. 24, 3137

(80) Chen, L.; Lyubimov, A. Y.; Brammer, L.; Vrielink, A.; Sampson, N. S. The Binding and Release of Oxygen and Hydrogen Peroxide Are Directed by a Hydrophobic Tunnel in Cholesterol Oxidase. *Biochemistry* **2008**, *47* (19), 5368–5377

(81) Liu, R.; Kang, Y.; Chen, L. Activation Mechanism of Human Soluble Guanylate Cyclase by Stimulators and Activators. *Nat. Commun.* **2021**, *12* (1), 5492

(82) Liu, R.; Kang, Y.; Chen, L. NO Binds to the Distal Site of Haem in the Fully Activated Soluble Guanylate Cyclase. *Nitric Oxide* **2023**, *134–135*, 17–22

(83) Wittenborn, E. C.; Thomas, W. C.; Houghton, K. A.; Wirachman, E. S.; Wu, Y.; Marletta, M. A. Role of the Coiled-Coil Domain in Allosteric Activity Regulation in Soluble Guanylate Cyclase. *Biochemistry* **2023**, *62* (10), 1568–1576

(84) Putnam, C. D.; Hammel, M.; Hura, G. L.; Tainer, J. A. X-Ray Solution Scattering (SAXS) Combined with Crystallography and Computation: Defining Accurate Macromolecular Structures, Conformations and Assemblies in Solution. *Q. Rev. Biophys.* **2007**, *40* (3), 191–285

(85) Skou, S.; Gillilan, R. E.; Ando, N. Synchrotron-Based Small-Angle X-Ray Scattering of Proteins in Solution. *Nat. Protoc.* **2014**, *9* (7), 1727–1739

(86) Morton, D. B. Invertebrates Yield a Plethora of Atypical Guanylyl Cyclases. *Mol. Neurobiol.* **2004**, *29* (2), 097–116

(87) Düner, M.; Lambertz, J.; Mügge, C.; Hemschemeier, A. The Soluble Guanylate Cyclase CYG12 Is Required for the Acclimation to Hypoxia and Trophic Regimes in *Chlamydomonas reinhardtii*. *The Plant Journal* **2018**, *93* (2), 311–337

(88) Reyes-Rivera, J.; Wu, Y.; Guthrie, B. G. H.; Marletta, M. A.; King, N.; Brunet, T. Nitric Oxide Signaling Controls Collective Contractions in a Colonial Choanoflagellate. *Curr. Biol.* **2022**, *32* (11), 2539-2547.e5

(89) Winger, J. A.; Derbyshire, E. R.; Lamers, M. H.; Marletta, M. A.; Kuriyan, J. The Crystal Structure of the Catalytic Domain of a Eukaryotic Guanylate Cyclase. *BMC Struct. Biol.* **2008**, *8* (1), 42

(90) Liping, Z.; Hongbo, S.; Xiaohua, L.; Zhaopu, L. Gene Regulation of Iron-Deficiency Responses Is Associated with Carbon Monoxide and Heme Oxydase 1 in *Chlamydomonas reinhardtii*. *PLoS One* **2013**, *8* (1), e53835

(91) Gray, J. M.; Karow, D. S.; Lu, H.; Chang, A. J.; Chang, J. S.; Ellis, R. E.; Marletta, M. A.; Bargmann, C. I. Oxygen Sensation and Social Feeding Mediated by a *C. elegans* Guanylate Cyclase Homologue. *Nature* **2004**, *430* (6997), 317–322

(92) Couto, A.; Oda, S.; Nikolaev, V. O.; Soltesz, Z.; de Bono, M. In Vivo Genetic Dissection of O_2 -Evoked CGMP Dynamics in a *Caenorhabditis elegans* Gas Sensor. *Proc. Natl. Acad. Sci. U S A* **2013**, *110* (35), 3301-3310

(93) Cheung, B. H. H.; Arellano-Carbajal, F.; Rybicki, I.; de Bono, M. Soluble Guanylate Cyclases Act in Neurons Exposed to the Body Fluid to Promote *C. elegans* Aggregation Behavior. *Curr. Biol.* **2004**, *14* (12), 1105–1111

(94) Morton, D. B. Atypical Soluble Guanylyl Cyclases in Drosophila Can Function as Molecular Oxygen Sensors. *J. Biol. Chem.* **2004**, *279* (49), 50651–50653

(95) Vermehren, A.; Langlais, K. K.; Morton, D. B. Oxygen-Sensitive Guanylyl Cyclases in Insects and Their Potential Roles in Oxygen Detection and in Feeding Behaviors. *J. Insect. Physiol.* **2006**, *52* (4), 340–348

(96) Langlais, K. K.; Stewart, J. A.; Morton, D. B. Preliminary Characterization of Two Atypical Soluble Guanylyl Cyclases in the Central and Peripheral Nervous System of *Drosophila melanogaster*. *J. Exp. Biol.* **2004**, *207* (13), 2323–2338

(97) Morton, D. B.; Langlais, K. K.; Stewart, J. A.; Vermehren, A. Comparison of the Properties of the Five Soluble Guanylyl Cyclase Subunits in *Drosophila melanogaster*. *Journal of Insect Science* **2005**, *5* (1)

(98) Huang, S. H.; Rio, D. C.; Marletta, M. A. Ligand Binding and Inhibition of an Oxygen-Sensitive Soluble Guanylate Cyclase, Gyc-88E, from *Drosophila*. *Biochemistry* **2007**, *46* (51), 15115–15122

(99) Giovine, M.; Pozzolini, M.; Favre, A.; Bavestrello, G.; Cerrano, C.; Ottaviani, F.; Chiarantini, L.; Cerasi, A.; Cangiotti, M.; Zocchi, E.; Scarfì, S.; Sarà, M.; Benatti, U. Heat Stress-Activated, Calcium-Dependent Nitric Oxide Synthase in Sponges. *Nitric Oxide* **2001**, *5* (5), 427–431

(100) Elliott, G. R. D.; Leys, S. P. Evidence for Glutamate, GABA and NO in Coordinating Behaviour in the Sponge, *Ephydatia muelleri* (Demospongiae, Spongillidae). *J. Exp. Biol.* **2010**, *213* (13), 2310–2321

(101) Say, T. E.; Degnan, S. M. Molecular and Behavioural Evidence That Interdependent Photo - and Chemosensory Systems Regulate Larval Settlement in a Marine Sponge. *Mol. Ecol.* **2020**, *29* (2), 247–261

(102) Moroz, L. L.; Meech, R. W.; Sweedler, J. V.; Mackie, G. O. Nitric Oxide Regulates Swimming in the Jellyfish *Aglantha digitale*. *Journal of Comparative Neurology* **2004**, *471* (1), 26–36

(103) Anctil, M.; Poulain, I.; Pelletier, C. Nitric Oxide Modulates Peristaltic Muscle Activity Associated with Fluid Circulation in the Sea Pansy *Renilla koellikeri*. *J. Exp. Biol.* **2005**, *208* (10), 2005–2017

(104) Moroz, L. L.; Romanova, D. Y.; Nikitin, M. A.; Sohn, D.; Kohn, A. B.; Neveu, E.; Varoqueaux, F.; Fasshauer, D. The Diversification and Lineage-Specific Expansion of Nitric Oxide Signaling in Placozoa: Insights in the Evolution of Gaseous Transmission. *Sci. Rep.* **2020**, *10* (1), 13020

(105) Bishop, C. D.; Brandhorst, B. P. NO/CGMP Signaling and HSP90 Activity Represses Metamorphosis in the Sea Urchin *Lytechinus pictus*. *Biol. Bull.* **2001**, *201* (3), 394–404

(106) Ribeiro, M.; Straub, V. A.; Schofield, M.; Picot, J.; Benjamin, P. R.; O'Shea, M.; Korneev, S. A. Characterization of NO-sensitive Guanylyl Cyclase: Expression in an Identified Interneuron Involved in NO–CGMP-dependent Memory Formation. *European Journal of Neuroscience* **2008**, *28* (6), 1157–1165

(107) Smith, B. C.; Marletta, M. A. Mechanisms of S-Nitrosothiol Formation and Selectivity in Nitric Oxide Signaling. *Curr. Opin. Chem. Biol.* **2012**, *16* (5–6), 498–506

(108) Edsinger, E.; Kieras, M.; Pirro, S. The Genome Sequences of 118 Taxonomically Diverse Eukaryotes of the Salish Sea. *Biodiversity Genomes* **2024**, https://doi.org/10.56179/001c.118307.

(109) Ellwanger, K.; Nickel, M. Neuroactive Substances Specifically Modulate Rhythmic Body Contractions in the Nerveless Metazoon *Tethya wilhelma* (Demospongiae, Porifera). *Front. Zool.* **2006**, *3* (1), 7

(110) Ueda, N.; Richards, G. S.; Degnan, B. M.; Kranz, A.; Adamska, M.; Croll, R. P.; Degnan, S. M. An Ancient Role for Nitric Oxide in Regulating the Animal Pelagobenthic Life Cycle: Evidence from a Marine Sponge. *Sci. Rep.* **2016**, *6* (1), 37546

(111) Musser, J. M.; Schippers, K. J.; Nickel, M.; Mizzon, G.; Kohn, A. B.; Pape, C.; Ronchi, P.; Papadopoulos, N.; Tarashansky, A. J.; Hammel, J. U.; Wolf, F.; Liang, C.; Hernández-Plaza, A.; Cantalapiedra, C. P.; Achim, K.; Schieber, N. L.; Pan, L.; Ruperti, F.; Francis, W. R.; Vargas, S.; Kling, S.; Renkert, M.; Polikarpov, M.; Bourenkov, G.; Feuda, R.; Gaspar, I.; Burkhardt, P.; Wang, B.; Bork, P.; Beck, M.; Schneider, T. R.; Kreshuk, A.; Wörheide, G.; Huerta-Cepas, J.; Schwab, Y.; Moroz, L. L.; Arendt, D. Profiling Cellular Diversity in Sponges Informs Animal Cell Type and Nervous System Evolution. *Science* **2021**, *374* (6568), 717–723

(112) Knoll, A. H. The Multiple Origins of Complex Multicellularity. *Annu. Rev. Earth Planet Sci.* **2011**, *39* (1), 217–239

(113) Goldstein, B.; King, N. The Future of Cell Biology: Emerging Model Organisms. *Trends Cell. Biol.* **2016**, *26* (11), 818–824

(114) Brunet, T.; King, N. The Origin of Animal Multicellularity and Cell Differentiation. *Dev Cell* **2017**, *43* (2), 124–140
(115) Brunet, T.; Larson, B. T.; Linden, T. A.; Vermeij, M. J. A.; McDonald, K.; King, N. Light-Regulated Collective Contractility in a Multicellular Choanoflagellate. *Science* **2019**, *366* (6463), 326–334

Chapter 2
Nitric oxide signaling controls collective contractions in a colonial choanoflagellate

Summary

Although signaling by the gaseous molecule nitric oxide (NO) regulates key physiological processes in animals, including contractility,[1–3] immunity,[4,5] development,[6–9] and locomotion,[10,11] the early evolution of animal NO signaling remains unclear. To reconstruct the role of NO in the animal stem lineage, we set out to study NO signaling in choanoflagellates, the closest living relatives of animals.[12] In animals, NO produced by the nitric oxide synthase (NOS) canonically signals through cGMP by activating soluble guanylate cyclases (sGCs).[13,14] We surveyed the distribution of the NO signaling pathway components across the diversity of choanoflagellates and found three species that express NOS (of either bacterial or eukaryotic origin), sGCs, and downstream genes previously shown to be involved in the NO/cGMP pathway. One of the species coexpressing sGCs and a bacterial-type NOS, *Choanoeca flexa*, forms multicellular sheets that undergo collective contractions controlled by cGMP.[15] We found that treatment with NO induces cGMP synthesis and contraction in *C. flexa*. Biochemical assays show that NO directly binds *C. flexa* sGC1 and stimulates its cyclase activity. The NO/cGMP pathway acts independently from other inducers of *C. flexa* contraction, including mechanical stimuli and heat, but sGC activity is required for contractions induced by light-todark transitions. The output of NO signaling in *C. flexa* – contractions resulting in a switch from feeding to swimming – resembles the effect of NO in sponges[1–3] and cnidarians,[10,16,17] where it interrupts feeding and activates contractility. These data provide insights into the biology of the first animals and the evolution of NO signaling.

This work was carried out in collaboration with Dr. Josean Reyes-Rivera and Dr. Thibaut Brunet in the lab of Prof. Nicole King and Dr. Benjamin G. H. Guthrie in the lab of Prof. Michael Marletta. J. R-R. and T. B. conceptualized and carried out cell biology assays detailed in this chapter; Y. W. and B. G. H. G. contributed the biochemical characterization of *Cf* sGC1 *in vitro*. This work led to the following publication: Reyes-Rivera, J., Wu, Y., Guthrie, B. G. H., Marletta, M. A., King, N., Brunet, N., *Current Biology* **2022,** 32, 2539-2547.

Methods

Culture of Choanoeca flexa

Colonies were cultured in 1% to 15% Cereal Grass Medium (CGM3) in artificial seawater (ASW). Polyxenic cultures (continuously passaged from a previously described environmental isolate[15]) were maintained at 22 °C under a light-dark cycle of 12:12 hours in a Caron low temperature incubator equipped with a lamp (Venoya Full Spectrum 150W Plant Growth LED) controlled by a programmable timer (Leviton VPT24^{-1}PZ Vizia). Polyxenic cultures used in most experiments were not light-sensitive, possibly due to progressive loss of bacterial diversity during serial passaging (as bacterially provided retinal is known to be required for photosensation in *C. flexa*[15]). Light-sensitive sheets used in photosensation experiments (Figures 2.4C and 2.4D) were thawed from stocks that had been frozen immediately after clonal isolation from a Curaçao isolate and cultured as described above.

Light microscopy—Imaging

Colonies were imaged in FluoroDishes (World Precision Instruments FD35-100) by differential interference contrast (DIC) microscopy using a 20x Zeiss objective mounted on a Zeiss Observer Z.1 with Hamamatsu Orca Flash 4.0 V2 CMOS camera (C1140-22CU).

Compound treatments and colony inversion assays

Small molecule inhibitor treatments and colony inversion assays were performed in 24-well plates (Fisher Scientific 09-761-146) containing 1 mL *C. flexa* culture per well. ODQ (pan-soluble guanylate cyclase inhibitor, BioVision 2051) was added 1 hour before behavioral assays. Addition of each small molecule compound was followed by a gentle swirl of the 24-well plate to ensure mixing. For each assay, all colonies visible within a well were counted (at least 30 colonies per biological replicate). All behavioral experiments were conducted under ambient light in the laboratory, unless indicated otherwise.

NO donor-induced inversion

The NO donors proliNONOate (Cayman Chemical Company 82145) and DEA-NONOate (Cayman Chemical Company 82100) were dissolved according to provider's instructions and stored as single-use aliquots at -80°C. Addition of NO donor proliNONOate induced inversion within 1-2 minutes. Prior to counting, colonies were fixed by addition of 16% ice-cold PFA in a 1:3 volumetric ratio, resulting in a final concentration of 4% PFA. Contracted and relaxed colonies were then manually counted by observation under a Leica DMIL LED transmitted light microscope.

Light-induced sheet inversion

After treatment with small molecule compounds, light-to-dark transitions were performed by manually switching off the light source of the DMIL LED microscope. The "light off" condition lasted for one minute before sheets were fixed and scored as described above.

Mechanically induced sheet inversion

3 mL of *C. flexa* culture were transferred to T12.5 culture flasks (Fisher Scientific 353107) and mechanically stimulated by vortexing on a Vortex Genie 2 (Scientific industries) on "Slow" setting for 5 seconds. Sheets were immediately fixed and scored as above.

Heat shock-induced sheet inversion

Colonies in 24-well plates were treated with inhibitors as described above and placed at the surface of a 37 °C warm bath for one minute. Sheets were immediately fixed and scored as above.

cGMP ELISA

For *in vivo* quantification of cGMP was performed with an ENZO Direct cGMP ELISA kit (ADI-900-014, 96 wells) as directed by the manufacturer. For each biological replicate, 90 mL of dense (>10^6 cells/mL) *C. flexa* culture was centrifuged for 5 minutes at 3000 x g and resuspended in 25 mL of ASW to wash the bacteria away. After the third wash, the cells were resuspended in 200 µL of ASW and split into one control (100 µL) and one treated sample (100 µL). The samples were lysed and quantified in parallel in each assay. Colonies from the "NO donor" group were

treated with 0.25 μM proliNONOate 5 minutes before lysis. Values were read on a SpectraMax M3 plate reader (Molecular Devices).

NO labeling, imaging, and image analysis

C. flexa cultures were transferred into 15 mL Falcon tubes and vortexed in "fast" setting on a Vortex Genie 2 for one minute to dissociate colonies into single cells. Cells were washed 3 times with artificial seawater (ASW) by centrifuging them for 5 minutes at 3000 x g and resuspending them in 25 mL of ASW. After the last wash, cells were resuspended in 1.5 mL ASW and transferred into a 1.5 mL Eppendorf tube. Cells were incubated in 10 μM DAF-FM (Invitrogen, D-23844) for 1 hour and rinsed twice with ASW to wash away the unincorporated dye. Cells were then transferred into a FluoroDish charged with a Corona surface treater and coated with poly-D-lysine, following a previously published protocol.[15] We let the cells adhere to the bottom of the dish for 30 minutes before imaging on a Z.1 Zeiss Imager with a Hamamatsu Orca Flash 4.0 V2 CMOS camera (C11440-22CU) and a 40x water immersion objective (C-Apochromat, 1.1 NA) for DIC and green epifluorescence imaging with a frame rate of 1 frame per minute. The NO donors (0.25 μM proliNONOate or 0.5 μM DEA-NONOate) were added 5 minutes after imaging began. We quantified intracellular fluorescence intensity using ImageJ. Change in fluorescence intensity was calculated by subtracting the fluorescence intensity at minute 1 from fluorescence intensity at minute 30.

Phylogenetic analysis and protein domain identification

We screened a selection of fully sequenced genomes for homologs of sGC and NOS with the following strategy: the protein sequences of Homo sapiens sGC-α1 and brain nitric oxide synthase (NOS1) were used as BLASTp queries against the NCBI database restricted to the following list of species:

Eukaryotes: *Homo sapiens* (Hsa), *Branchiostoma floridae* (Bfl), *Drosophila melanogaster* (Dme), *Capitella teleta* (Cte), *Nematostella vectensis* (Nve), *Amphimedon queeslandica* (Amq), *Mnemiopsis leidyi, Trichoplax adhaerens* (Tadh), *Salpingoeca rosetta* (Sro), *Capsaspora owczarzaki, Sphaereoforma arctica* (Sphac), *Abeoforma whisleri, Creolimax fragrantissima, Pirum gemmata, Aspergillus oryzae* (Asory), *Jimgerdemannia flammicorona* (Jifla), *Rhizoctonia solani* (Rhiso), *Pterula gracilis* (Ptegra), *Schizosaccharomyces pombe, Tuber melanosporum, Cryptococcus neoformans, Ustilago maydis, Cryptococcus neoformans, Ustilago maydis, Rhizopus oryzae, Allomyces macrogynus, Batrachochytrium dendrobatidis, Spizellomyces punctatus, Thecamonas trahens, Dictyostelium discoideum, Polysphondylium pallidum, Entamoeba histolytica, Arabidopsis thaliana, Selaginella moellendorffii, Physcomitrella patens, Chlamydomonas reinhardtii, Volvox carteri* (Vcar), *Chlorella variabilis* (Chl), *Ostreococcus tauri* (Ostau), *Ectocarpus siliculosus, Phaeodactylum tricornutum, Thalassiossira pseudonana, Phytophthora infestans*, *Toxoplasma gondii, Tetrahymena thermophila, Perkinsus marinus, Guillardia theta, Naegleria gruberi* (Ngru), *Trypanosoma cruzi, Leishmania major, Trichomonas vaginalis, Giardia lamblia, Bigelowiella natans, Emiliania huxleyi*

Archaea: *Nanoarchaeum equitans, Ignicoccus islandicus, Natronolimnobius baerhuensis, Halorientalis regularis, Halostagnicola kamekurae, Halalkalicoccus subterraneus,*

Halobiforma nitratireducens (Halob), *Natronobacterium gregoryi* (Natr), *Haloplanus natans, Halovenus aranensis* (Halar), *Halonotius pteroides* (Halo)

Bacteria: *Actinocrispum wychmicini* (Actin), *Kibdelosporangium aridum* (Kibd), *Crossiella equi* (Cross), *Lentzea xinjiangensis* (Lentz)Nocardioides *speluncae* (Noc), *Saccharopolyspora spinosa* (Sspi), *S ynechococcus sp.* PCC 7335 (Syn), *Nostoc cycadae* (Nocyc), *Anabaenopsis circularis* (Ancir), *Planktothrix paucivesiculata* (Plank), *Crinalium epipsammum* (Crinep), *Spirosoma radiotolerans* (Spiro), *Roseinatronobacter monicus* (Rose)

Additional BLASTp searches were conducted against a published dataset of 19 choanoflagellate transcriptomes,[18] the *C. flexa* transcriptome,[15] the *Mnemiopsis leidyi* genome (https://research.nhgri.nih.gov/mnemiopsis/sequenceserver/) and the *Ministeria vibrans* transcriptome[19] (and courtesy of Daniel J. Richter). The *C. flexa* NOS and sGC predicted protein sequences were deposited onto NCBI with the following accession numbers: GenBank: ON075806 (for *Cf* NOS), GenBank: ON075810 (for *Cf* sGC1), GenBank: ON075809 (for *Cf* sGC2), GenBank: ON075808 (for *Cf* sGC3), and GenBank: ON075807 (for *Cf* sGC4).

Domain architectures were predicted using the CD-search tool from NCBI. For phylogenetic reconstructions, sequences were aligned using Clustal implemented in Geneious Prime (2021 version). The NOS sequence alignment was manually trimmed to be restricted to the oxidase domain and the sGC alignment was trimmed using gBlocks with minimally stringent parameters. Phylogenetic trees were reconstructed using PhyML and BMGE implemented on http://phylogeny.lirmm.fr/phylo_cgi/index.cgi.[70] Trees were visualized using iTOL (https://itol.embl.de/)[20] and further edited in Adobe Illustrator 2021. Species silhouettes were added from PhyloPic (http://phylopic.org/).

Construction of expression plasmid

First-strand *C. flexa* cDNA (extracted as in[15]) was used as the template for cloning *Cf* sGC1. Forward and reverse primers were designed against 5' and 3' ends of the target transcript (transcript name: TRINITY_DN6618_c0_g1_i1 in the published transcriptome[15]). Forward: TAAGAAGGAGATATACCATGTATGGCTTGGTGCACGAAGC; reverse: TAATGGTGATGATGGTGATGAACTATAGTCTGCTTGCCAACG. Underlined portions anneal to the sequence template. The PCR product was inserted into a pET28b vector using Gibson assembly, and the cloning product was verified by sequencing (UC Berkeley sequencing facility).

Protein expression and purification

pET_*Cf* sGC1 was transformed into *E. coli* BL21star (DE3) cells co-expressing the chaperone GroEL/ES from the pGro7 plasmid (Takara Biosciences). After overnight incubation at 37 °C in LB Miller media supplemented with 50 μg/mL kanamycin, 20 μg/mL chloramphenicol and 500 μM iron (III) chloride, cells were subcultured 1:200 into TB media supplemented with 50 μg/mL kanamycin, 20 μg/mL chloramphenicol, 500 μM iron (III) chloride, 0.5 mg/mL L-arabinose, and 2 mg/mL glucose, grown at 37 °C. Once cell density reached OD_{600} = 0.6, 1 mM 5-aminolevulinic acid was added to the culture, and culturing temperature was lowered to 18 °C. After 15 minutes of incubation, protein production was induced

by addition of 500 μM isopropyl β-D-1-thiogalactopyranoside and cultures were incubated for an additional 18 hours. Cell culture was harvested by centrifuging at 4200 x g for 25 min. Cells were collected, flash frozen in liquid nitrogen, and stored at -80 °C until purification.

All protein purification steps were done at 4 °C unless otherwise noted. Cells were resuspended in equal volume of buffer A (50 mM sodium phosphate, 150 mM NaCl, 5 mM imidazole, 5% glycerol, pH 8.0) supplemented with 110 mM benzamidine, 0.4 mM AEBSF, and 0.3 mg/mL DNAse. Cell resuspension was lysed using an Avestin EmulsiFlex-C5 homogenizer. Cell lysate was collected and clarified by centrifugation at 32,913 g for 55 min, and the supernatant was collected and loaded onto a His60 Ni Superflow gravity column (Takara Bio). The column was washed twice, first with 10 CV buffer A, and then with 10 CV of a 9:1 mixture of buffer A and buffer B (50 mM sodium phosphate, 150 mM NaCl, 400 mM imidazole, 5% glycerol, pH 8.0). Protein was eluted with 5 CV buffer B in 1 mL fractions. Fractions with heme absorbance were pooled and concentrated using a 50 kDa cutoff spin concentrator and supplemented with 5 mM DTT and 1 mM EDTA for overnight storage. Subsequently, protein was passed over a POROS HQ2 anion exchange column (Applied Biosystems). After loading, the column was washed with 5 CV of buffer C (25 mM triethanolamine, 25 mM NaCl, 5 mM DTT, pH 7.4) and developed over a gradient of 100 mM – 300 mM NaCl over 17 CV. The protein absorption spectrum was measured using a Nanodrop 2000 microvolume spectrophotometer (ThermoFisher Scientific). Subsequently, protein was aliquoted, flash frozen in liquid nitrogen, and stored at -80 °C for future use.

Analytical size exclusion chromatography

Purified *Cf* sGC1 and the protein standard mixture were injected onto a Superdex 200 Increase 10/300 GL column (GE healthcare) equilibrated with buffer F (50 mM triethanolamine, 150 mM NaCl, 5 mM DTT, 5% glycerol). Protein elution was monitored by UV absorbance at 280 nm.

Gas ligand binding of bacterial-produced Cf sGC1

Cf sGC1 was handled in an argon-filled glove bag. Protein-bound heme was reduced by adding sodium dithionite (~500-fold excess over the protein conc.). Excess dithionite was removed by gel filtration of the protein into Buffer E (50 mM HEPES, 150 mM NaCl, 5% glycerol, pH 7.4) using a pre-equilibrated Zeba spin desalting column. A ferrous, ligand-free UV-vis absorption spectrum was collected on a Cary 300 UV-vis Spectrophotometer. Fe(II)-NO bound *Cf* sGC1 was generated by adding the NO-releasing molecule proliNONOate (~10-fold excess) to the protein sample, and Fe(II)-CO bound *Cf* sGC1 was generated by adding CO-sparged Buffer E to the protein sample before collecting a spectrum.

Extinction coefficient of Cf sGC1

The extinction coefficient of the Soret maximum of reduced *Cf* sGC1 was measured using two assays performed in tandem: heme concentration in a sample of *Cf* sGC1 was determined using pyridine hemochromagen assay, and the heme Soret absorption of the protein sample was measured using UV-vis spectroscopy as described above. The pyridine hemochromagen assay was carried out following a reported protocol.[21] Briefly, a reduced protein sample with a known Soret band absorbance was diluted 5-fold in Buffer E, and then further diluted 2-fold in Solution I (0.2 M NaOH, 40% pyridine, 500 μM potassium ferricyanide) to yield the oxidized pyridine

hemochromagen. An aliquot (10 μL) of Solution III (0.5 M sodium dithionite, 0.5 mM NaOH) was then added to the oxidized pyridine hemochromagen sample to yield the reduced pyridine hemochromagen. The UV-vis absorption spectrum of the reduced pyridine hemochromagen was measured on a Cary 300 UV-vis spectrophotometer. Absorption at 557 nm (ε = 34,700 M^{-1} cm^{-1}) was used to calculate the heme concentration in the pre-dilution sample, and the extinction coefficient of the heme cofactor of *Cf* sGC1 was calculated by dividing the reduced heme absorption by the heme concentration.

Activity assays and quantification

Specific activity for *Cf* sGC1 was measured by quantifying the amount of cGMP produced in duplicate end-point activity assays, done in biological triplicate. *Cf* sGC1 from previously frozen aliquots was thawed and reduced in an anaerobic chamber as described above. The reduced protein was used without further treatment for the basal (unliganded) activity assay. A UV-vis spectrum for the unliganded sample was obtained using a Nanodrop 2000 microvolume spectrophotometer, and the Soret maximum was used to quantify heme-bound protein concentration (ε_{428} = 101,000 M^{-1} cm^{-1}). Only protein with a Soret:280 ratio > 1 was used for activity assays. To the remaining reduced protein, proliNONOate was added to a final concentration of 400 μM by the addition of 1 μL of a stock solution of 10 mM proliNONOate in 10 mM NaOH, and the protein sample was incubated at 4 °C for 5 min to yield the xsNO sample. The xsNO-bound UV-vis spectrum was then collected. The protein concentration of the xsNO sample was assumed to be the same as the unliganded sample. The xsNO sample was then buffer exchanged by gel filtration using a Zeba spin column into buffer E to yield the 1-NO sample, and the UV-vis spectrum was collected. The concentration of the 1-NO sample was calculated based on the NO-bound Soret absorbance at 399 nm compared to that of the xsNO sample. Activity assays were carried out at 25 °C in Buffer E supplemented with 5 mM DTT and 3 mM $MgCl_2$ with *Cf* sGC1 concentration at 40 nM. To obtain the xsNO state, 70 μM proliNONOate was added. The reaction was initiated by addition of 1.5 mM GTP, and timepoints were quenched by diluting the reaction mixture 1:4 to a solution of 125 mM zinc acetate, and pH adjusted by adding equal volume of 125 mM sodium carbonate. Assay samples were stored at -80 °C until analyzed. Quenched assay samples were thawed at room temperature, centrifuged at 21,130 x g at 4 °C to remove zinc precipitate. Supernatant was collected and diluted 250-fold for the analysis. cGMP content of each assay sample was quantified in duplicate using an enzyme-linked immunosorbent assay (Enzo Life Sciences) following the manufacturer's protocol. Initial rate of the reaction was calculated using the linear phase of the time course, where <10% of substrate has been depleted.

Quantification and statistical analysis

Information about the quantification and statistical details of experiments can be found in the corresponding figure legends. Statistical tests and graphs were produced using Prism 9.0.0.

Results

C. flexa encodes both NOS and sGC

C. flexa is a colonial choanoflagellate that forms concave sheets capable of global inversion of their curvature through collective contractility.[15] *C. flexa* inversion mediates a trade-off between feeding and swimming: relaxed colonies (with their flagella pointing inside) are slow swimmers and efficient feeders, whereas contracted colonies (with their flagella pointing outside) are

inefficient feeders but fast swimmers (Figure 2.1A).[15] Inversion has been shown to be induced by light-to-dark transitions through the inactivation of a rhodopsin phosphodiesterase (Rho-PDE) and accumulation of cyclic guanosine monophosphate (cGMP) (Figure 2.1B).[15] The involvement of cGMP in collective contraction in *C. flexa* and the connection between nitric oxide (NO)/cGMP signaling and tissue contraction in nonbilaterian animals[22] led us to investigate whether NO signaling might exist and regulate sheet inversion in *C. flexa* (Figure 2.1B).

By examining the genomes of two choanoflagellates (*Monosiga brevicollis*[23] and *Salpingoeca rosetta*[24]) and the transcriptomes of *C. flexa*[15] and 19 other choanoflagellates,[18] we found that five choanoflagellate species encode nitric oxide synthase (NOS) homologs and that nine choanoflagellate species encode soluble guanylate cyclase (sGC) homologs (Figure 2.2A). Of these, three species—*Salpingoeca infusionum*, *Choanoeca perplexa*, and *C. flexa*—express both NOS and sGCs, suggesting that some aspect of the physiology of these organisms might be regulated by NO/cGMP signaling.

Like metazoan NOS genes, choanoflagellate NOS genes encode the canonical oxygenase and reductase domains (Figure 2.2B). Phylogenetic analysis revealed that one choanoflagellate NOS gene, from *S. infusionum*, is closely related to those from metazoans and fungi (Figure A.1A). In contrast, all other choanoflagellate NOSs (including those from *C. flexa*) cluster with cyanobacterial NOSs, suggesting that they might have been acquired through horizontal gene transfer from a cyanobacterial ancestor early in the evolution of choanoflagellates (Figure A.1A). Current uncertainties surrounding the choanoflagellate phylogeny[25,26] make it difficult to pinpoint exactly when this horizontal gene transfer event occurred and whether it followed or preceded the loss of the ancestral eukaryotic NOS in most or all choanoflagellates. These choanoflagellate NOS genes, like those of cyanobacteria, differ from metazoan NOS, in which they encode an upstream globin domain with unknown function and lack the calmodulin-binding domain that mediates regulation of metazoan NOSs by Ca^{2+}, suggesting that calcium signaling does not regulate NO synthesis in *C. flexa*. The *C. flexa* transcriptome also encodes complete biosynthetic pathways for the NOS cofactors tetrahydrobiopterin (H_4B), flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD), and nicotinamide dinucleotide phosphate (NADPH) (Figure A.2A) as well as downstream genes in the NO/cGMP signaling pathway: cGMP-dependent kinase (PKG), cGMP-gated ion channels (CNG), and cGMP-dependent PDE (PDEG) (Figure A.2B), providing additional evidence that *C. flexa* employs NO signaling as part of its physiology.

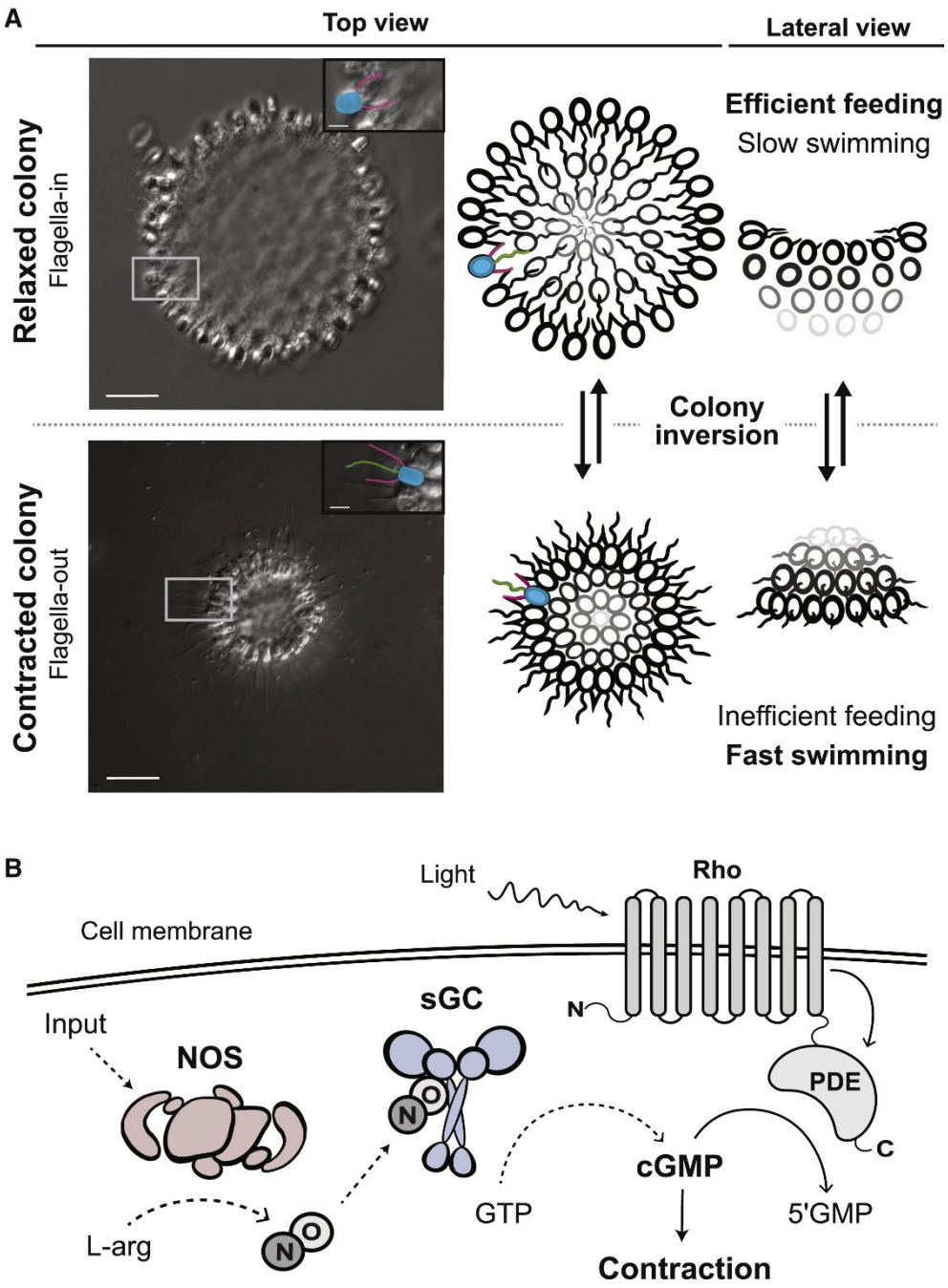


Figure 2.1. The inversion behavior of *C. flexa*, its control by light, and hypothesized control by NO. (A) Micrograph (left) and schematic depiction (right) of *C. flexa* inversion behavior. Cells are linked by their collars and form a cup-shaped monolayer or sheet; scale bars, 15 µM. Relaxed colonies (top half) have their flagella pointing toward the inside of the colony and are efficient feeders and slow swimmers. (Flagella from relaxed colony are not in focus in the micrograph; see schematic for flagellum orientation.) Contracted colonies (bottom half) have their flagella pointing toward the outside of the colony and are inefficient feeders but fast swimmers. Insets: pseudocolors highlight the characteristic morphological features of choanoflagellates: flagella (green), cell body (blue), and microvilli (magenta). Inset scale bars, 5 µm. (B) *C. flexa* colony inversion is controlled by light-to-dark transitions, mediated by a rhodopsin-phosphodiesterase fusion protein (Rho-PDE) upstream cGMP signaling. In the presence of light, Rho-PDE is active and constantly converting cGMP into 5′GMP. We hypothesized that NO/cGMP signaling might also be able to induce inversion. Mechanisms tested in this paper are indicated with dashed lines: a primary

input activates the nitric oxide synthase (NOS), which converts L-arginine into NO and L-citrulline. NO diffuses away and activates soluble guanylate cyclase (sGC) that convert GTP into cGMP, causing colony contraction.

Nearly all sGC genes from choanoflagellates, including *C. flexa*, encode the canonical domains observed in animal sGCs: the heme NO-/O_2-binding domain (HNOB or H-NOX), the HNOB-associated domain (HNOBA), and the C-terminal catalytic domain (guanylate cyclase) (Figure 2.2C).[27] Phylogenetic analysis revealed that all animal and most choanoflagellate sGCs (including those of *C. flexa*) evolved from a single ancestral sGC found in the last common ancestor of choanoflagellates and metazoans that diversified separately into multiple paralogs in these two lineages (Figure A.1B). In contrast, the predicted sGC from one choanoflagellate species, *S. helianthica*, more closely resembles the sGCs of chlorophyte algae (which are the only protist group previously known to encode sGC proteins with a metazoan-like domain architecture; Figure A.1B).[28]

Importantly, not all animal sGCs are regulated by NO: in *Drosophila melanogaster* and *Caenorhabditis elegans*, the so-called ‘‘atypical sGCs’’ preferentially bind soluble O_2 instead of NO and are thought to be involved in the regulation of feeding by oxygen concentration.[29–31] Discrimination between NO and O_2 is mediated by the presence of a hydrogen-bonding network, where a distal pocket tyrosine residue is critical for stabilizing the heme-O_2 complex.[32] We generated preferential binding predictions based on this motif and found that both NO- and O_2-preferential binding sGCs are widely distributed among choanoflagellates with no obvious pattern (Figures 2.2A and A.2D).[32,33] Interestingly, predicted NO-selective sGCs are present in choanoflagellate species in which NOS was not detected, suggesting that these species might detect NO from an exogenous source (i.e., environmental bacteria or other protists), might possess an alternative NO-producing mechanism, or might encode an NOS that was not detected in the transcriptome. In *C. flexa*, one of the four sGC transcripts was predicted to be selective for NO and was named *Cf* sGC1 (Figure A.2D). The other two choanoflagellate species found to possess both an NOS and sGCs, *C. perplexa* and *S. infusionum*, were also predicted to encode at least one NO-sensitive sGC (Figure 2.2A).

NO/cGMP signaling controls colony contraction in C. flexa

To test whether NO signaling regulates collective contractions in *C. flexa*, we treated *C. flexa* cultures with several NO donors, compounds capable of releasing NO in solution.[34] We found that treatment of *C. flexa* with the NO donors proliNONOate and DEA-NONOate led to an increase in intracellular NO as detected by the NO-sensitive fluorescent probe 4-Amino-5-Methylamino-2’,7’-Difluorofluorescein (DAF-FM), demonstrating that they could be effective reagents for studying NO signaling in *C. flexa in vivo* (Figures 2.3A and A.3A–A.3C). Treatment of *C. flexa* with proliNONOate (Figures 2.3B and 2.3C) induced colony contraction within 1-2 min (although the inversion process itself only lasted a few seconds, as previously described for darkness-induced inversion;[15] see Video S1). The inversion response of *C. flexa* to proliNONOate was concentration dependent (Figure 2.3D), reaching a plateau of nearly 100% inversion at a concentration of 0.1 mM. As a negative control, *C. flexa* did not invert in response to proline, the molecular backbone of proliNONOate and the end product of NO release (Figure 2.3D). Moreover, treatment with two other NO donors, DEA-NONOate or NOC-12, also induced inversion (Figure A.3D). Taken together, these results indicate that NO is sufficient to induce colony contraction.

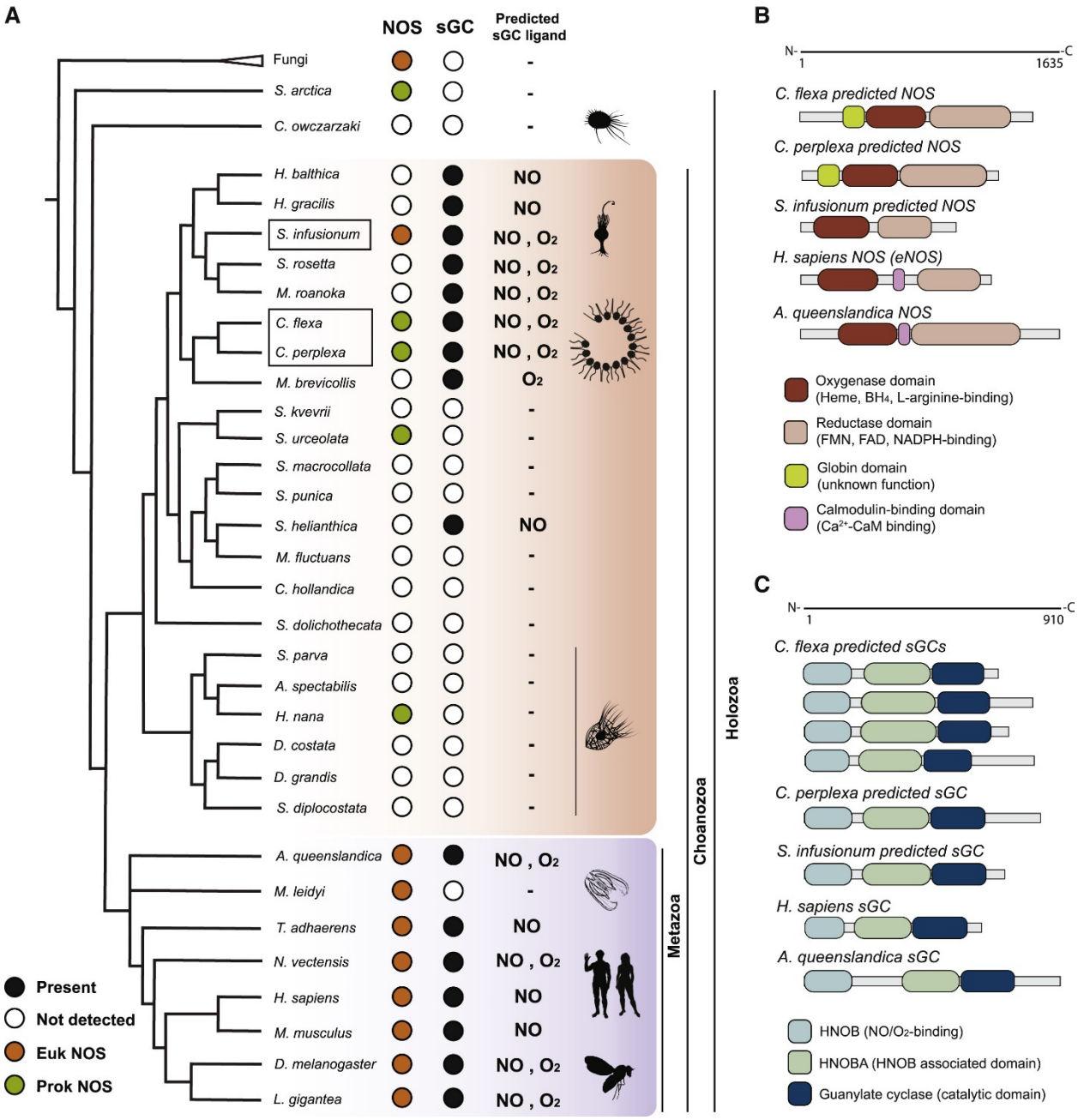


Figure 2.2. NO synthase (NOS) and soluble guanylate cyclases (sGCs) predicted to bind either NO or O_2 are broadly distributed across choanozoans, and all three are present in *C. flexa*. (A) Phylogenetic distribution of NOS and sGC across opisthokonts. *C. flexa*, its sister species *C. perplexa*, and *S. infusionum* encode both NOS and NO-sensitive sGCs, as do animals. (B) Choanoflagellate NOSs have the metazoan canonical oxygenase and reductase domain but lack the calcium-calmodulin binding domain. *C. flexa* and other choanoflagellate NOS encode an upstream globin domain with unknown function, also observed in cyanobacterial NOS. (C) *C. flexa* sGCs have the same domain architecture as animal sGCs. See Figure A.1 for phylogenetic analysis and Figure A.2 for phylogenetic distribution of NO/cGMP downstream components.

We next investigated whether NO triggers cGMP synthesis in *C. flexa*. We found that a pan-inhibitor of sGCs, oxadiazolo-quinoxalin-1-one (ODQ),[35] had no detectable effect on the percentage of contracted colonies 5 min after treatment with NO relative to cultures not treated

with ODQ. However, a difference between the two conditions became evident 10 min after NO treatment, with nearly all ODQ-treated colonies relaxing into the uncontracted form, although all ODQ-untreated colonies remained contracted at least twice as long (Figure 2.3E). This premature relaxation of ODQ-treated colonies was concentration dependent (Figure 2.3F).

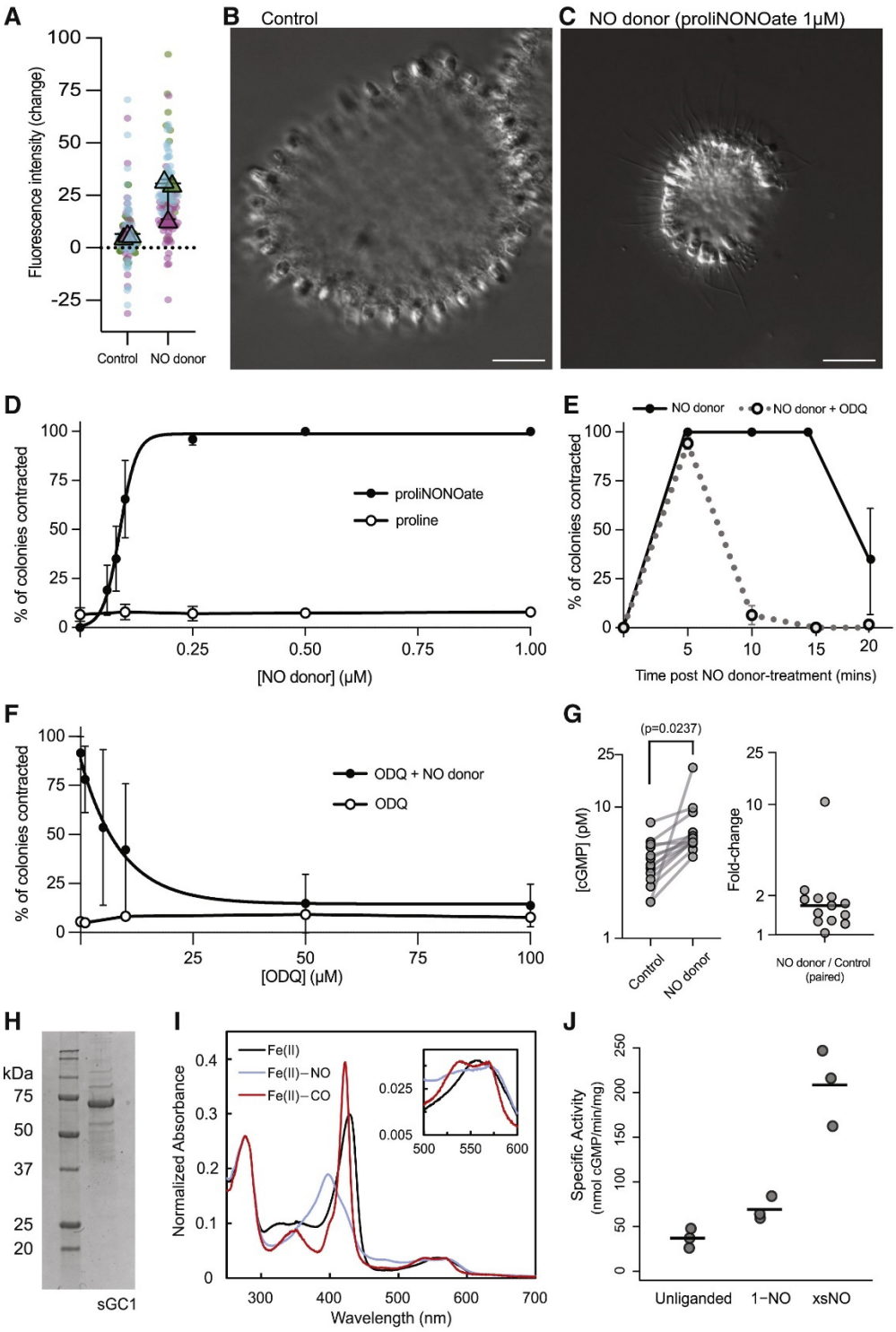


Figure 2.3. NO induces sustained colony contraction in *C. flexa* and activates *Cf* sGC1. (A) NO release by proliNONOate was visualized by loading cells with the NO-sensitive fluorescent probe DAF-FM and measuring the change in intracellular fluorescence intensity over 30 min. Cells treated with NO donor showed a higher increase in fluorescence. Graph is a SuperPlot[31] of n = 3 biological replicates per condition, with sample sizes N = 27, 33, and 47 for the three control groups and N = 26, 57, and 42 for the three treated groups. p = 0.0331 by an unpaired t test. See Figures A.3A–A.3C for a similar analysis using the NO donor DEA-NONOate. (B) and (C) NO induces colony contraction. Although colonies treated with a negative control compound (proline) remained relaxed, colonies treated with 0.25 μM of the NO donor proliNONOate contracted within 1–2 min. (D) The NO donor proliNONOate induces contractions within ∼1–2 min in a dose-dependent manner. A closely related molecule incapable of releasing NO (proline) had no effect over the same concentration range. Figure A.3D shows NO-induced contractions using different NO donors. Error bars indicate standard deviations. (E) Inhibition of sGCs with 50 μM ODQ did not abolish NO-induced contraction at early time points but greatly reduced its duration. ODQ-treated colonies were contracted within 5 min post-treatment with NO donor but relaxed soon after, whereas untreated colonies remain contracted for at least 5 more min. Error bars indicate standard deviations. (F) Inhibition of sustained contraction by ODQ is dose dependent. Colonies were incubated with different concentrations of ODQ for 1 h before treatment with 0.25 μM of the NO donor proliNONOate. We quantified the percentage of contracted colonies 10 min after NO treatment. In (D)–(F), each point represents the mean value of n = 3 biological replicates with at least 30 colonies scored per biological replicate. Error bars are standard deviations. (G) Treatment of cells with a NO donor increased intracellular cGMP concentration almost 2-fold as quantified by ELISA. N = 13 pairs of control/treated samples, p = 0.024 by a paired t test. (H)–(J) Purification, ligand binding properties, and specific activity of *Cf* sGC1. (H) Coomassie-stained SDS-PAGE gel of recombinant *Cf* sGC1 expressed in *E. coli* and purified. Band in lane "sGC1" represents *Cf* sGC1 with a monomeric molecular weight of 75.7 kDa. Left lane, molecular weight ladder. (I) UV-visible absorption spectra of *Cf* sGC1 under unliganded (black), NO-bound (blue), and CO-bound (red) conditions. Soret maxima: NO-bound: 429 nm; CO-bound: 423 nm; NO-bound: 399 nm. Inset, α/β bands show increased splitting upon ligand binding. (J) Specific activity of *Cf* sGC1 under unliganded, equimolar quantities of NO (1-NO) and excess NO (xsNO) conditions. Initial rates were measured from activity assays performed at 25°C, pH 7.5 with 1.5 mM Mg^{2+}-GTP substrate and 40 nM enzyme. Horizontal bars represent mean of three biological replicates.

These results suggest that the maintenance of NO-induced colony contraction requires sGC activity. However, the pan-sGC inhibitor ODQ does not allow us to discriminate between NO- and O_2-sensitive sGCs. If NO does indeed signal through the sGC/cGMP pathway, we predicted that treatment of colonies with an NO donor should increase intracellular cGMP concentration. We found that NO-treated cells contained consistently higher intracellular cGMP levels than untreated cells (2-fold increase on average, p = 0.0237; Figure 2.3G). Importantly, prior work has demonstrated that treatment of *C. flexa* with a cell-permeant form of cGMP (8-Br-cGMP) is sufficient to induce inversion.[15] These results further support the hypothesis that NO activates sGCs and stimulates the production of cGMP, which might contribute to maintaining *C. flexa* colony contraction.

C. flexa sGC1 is an NO-selective and catalytically active component of the C. flexa NO/cGMP signaling pathway

Of the four sGCs expressed by *C. flexa*, only one (*Cf* sGC1) was predicted to be selective for NO (Figure 2.1A). To directly test the existence of a canonical NO/sGC/cGMP pathway in *C. flexa*, we characterized ligand binding and activity of *Cf* sGC1 that had been heterologously expressed in and purified from *E. coli* (Figure 2.3H). The ligand specificity of sGCs can be determined by comparing the ''Soret peak'' maxima of the ultraviolet-visible (UV-vis) absorption after incubation with different gases, including NO, O_2, or CO. The Fe(II), unliganded form of *Cf* sGC1 exhibited a Soret peak at 429 nm and a single broad plateau in the α/β region (500–600 nm), consistent with that of an NO-selective sGC (Figure 2.3I; Table A.1).[33] In the presence of NO, the Soret peak shifted to 399 nm with two peaks in the α/β region, as characteristic for a 5-coordinate, high-spin NO-heme complex (Figure 2.3I; Table A.1).[33] On the other hand, when exposed to

atmospheric oxygen in the absence of NO, the UV-vis absorption spectrum of *Cf* sGC1 did not change (Figures 2.3I and A.4B), consistent with our prediction that *Cf* sGC1 is an NO-selective sGC and does not bind O_2. Moreover, like previously characterized animal sGCs, *Cf* sGC1 also binds CO to form a 6-coordinate, low-spin CO-heme complex, evidenced by a Soret band maximum of 424 nm with two peaks in the α/β region (Figure 2.3I; Table A.1).[28] Thus, the UV-vis spectroscopy results indicate that *Cf* sGC1 binds diatomic gas ligands in a manner that resembles other well-characterized NO-selective sGCs. Finally, size exclusion chromatography showed that *Cf* sGC1 formed a homodimer (Figure A.4A), similar to the dimeric structure of animal sGCs. In animals, NO-selective sGCs display three levels of activity: (1) in the absence of NO, the protein has a low basal guanylate cyclase activity; (2) when one NO molecule is bound at the heme moiety, the protein is partially activated (to several-fold the basal activity); and (3) in the presence of excess NO, the protein reaches maximal activation.[36] To characterize the enzymatic activity of *Cf* sGC1, we measured cGMP production by purified *Cf* sGC1 under unliganded, equimolar NO, and excess NO conditions using an endpoint activity assay. We found that *Cf* sGC1 has an activity profile similar to that of animal sGCs, with a 2-fold increase in activity under equimolar NO concentration and 6- fold increase under excess NO (Figure 2.3J). Overall, *Cf* sGC1 exhibits similar ligand binding properties and NO-stimulated activity profile to animal NO-specific sGCs.[37] These results further support the existence of NO/cGMP signaling in *C. flexa* and is consistent with it being mediated (at least in part) by *Cf* sGC1.

NO-cGMP signaling acts independently from most other inducers of colony contraction

In animals, NO signaling can be induced by a broad range of stimuli, which include chemical signals (for example, acetylcholine in mammalian blood vessels[38]), mechanical cues (for example, shear stress in blood vessels[39]), or heat shocks.[40–43] Interestingly, collar contractions in choanoflagellates can often be induced by mechanical stimuli, such as flow and touch.[44–47] We thus set out to test whether NO signaling in *C. flexa* responds to or intersects with environmental inducers of inversion.

We observed that *C. flexa* colonies invert in a matter of seconds in response to agitation of culture flasks (which presumably combines the effect of flow and shocks with other colonies or the walls of the flask) and to heat shocks (Figures 2.4A and 2.4B). To test whether mechanically or heat-induced contraction requires NO/cGMP signaling, we incubated the colony cultures with the pan-sGC inhibitor ODQ and exposed them to either of the two different stressors. We found that the inhibition of sGCs did not abolish mechanically induced or heat shock-induced contraction (Figures 2.4A and 2.4B). Taken together, these findings suggest that the mechanosensitive and thermosensitive pathways that induce inversion in *C. flexa* are independent of sGCs and hint at complex behavioral regulation in this choanoflagellate.

Previous work has shown that *C. flexa* colonies invert in response to light-to-dark transitions that they detect through a Rho-cGMP pathway.[15] In the presence of light, a Rho-PDE hydrolyzes cGMP into 5'-GMP, thus preventing cGMP signaling. In darkness, the Rho-PDE is inactivated, which allows cGMP to accumulate and trigger colony inversion (Figure 2.1B). Interestingly, this pathway requires the presence of cGMP, which is presumably synthesized by either particulate (i.e., membrane bound) or soluble (i.e., cytosolic) guanylate cyclases,[48] both of which are predicted to be encoded by the *C. flexa* transcriptome (Figure A.2C). A third family of

guanylate cyclases – NIT-GCs, recently discovered in animals – could not be detected in choanoflagellates.[49]

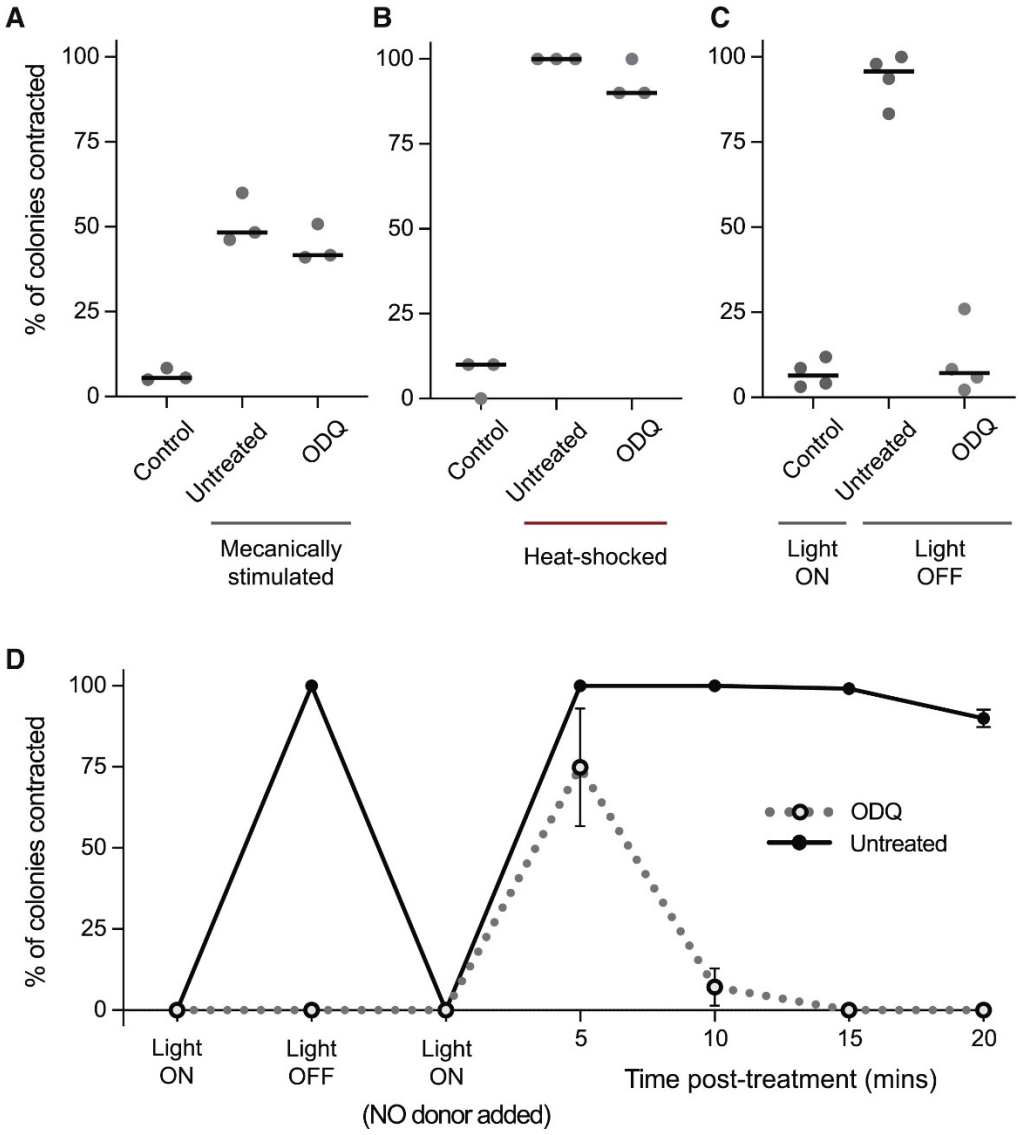


Figure 2.4. NO/cGMP acts independently of most other inducers of contraction in *C. flexa* (A) and (B) Mechanical stimuli and heat shock induce *C. flexa* colony contraction. Inhibition of sGC (50 μM ODQ) did not have an effect on mechanically or heat shock-induced contractions. (C) Inhibition of sGC (50 μM ODQ) abolished darkness-induced inversion ($p < 0.0001$ by an unpaired t test). (D) Colonies treated with ODQ did not respond to on/off changes in light but briefly contracted when exposed to NO donor, consistent with earlier results (Figures 2.3D–2.3F). Ten min after treatment, ODQ-treated colonies were relaxed, whereas untreated colonies remained contracted for a longer period. In (A)-(D), each point is the average of n = 3 (A) and (B) or n = 4 (C) and (D) biological replicates with at least N = 30 colonies scored per biological replicate. Error bars are standard deviations.

We next set out to answer whether sGCs are necessary for synthesizing the cGMP required for phototransduction. To address this, we treated light-sensitive colonies with ODQ and found that this entirely abolished darkness-induced inversion (Figure 2.4C). These results suggest that sGCs (either NO dependent or O_2 dependent) are responsible for synthesizing baseline levels of

cGMP that are then used during phototransduction. Even though ODQ-treated light-sensitive colonies did not invert in response to darkness, we confirmed that they could still respond to NO by undergoing brief contractions (which were sustained for a much shorter time than in controls; Figure 2.4D), consistently with earlier results (Figure 2.3E).

Discussion

Here, we report the presence of NOS, sGCs, and downstream components of NO-cGMP signaling in three choanoflagellate species, at least two of which (*C. flexa* and *C. perplexa*) are capable of collective contractions.[15,50] To our knowledge, this is the first observation of both NOS and sGCs in a nonanimal model. We found that NO causes sustained colony contraction in *C. flexa* and an increase in cGMP concentration in live cells, and that inhibition of sGCs (and thereby reduction in cGMP concentration) accelerated colony relaxation. Moreover, *in vitro* experiments confirmed that NO directly binds *Cf* sGC1, which it activates with a two-step profile in response to different NO concentration, as in animal sGCs.[36]

The observation that colonies treated with the sGC inhibitor initially contracted in response to NO at levels matching untreated colonies, only to relax much more quickly, was unexpected. We hypothesize that NO-induced contractions are mediated through at least two different pathways: a slow pathway (described above) that maintains contraction and requires sGC/cGMP and an (unidentified) fast pathway independent of sGC/cGMP. Moreover, treatment of light-sensitive colonies with the sGC inhibitor abolished darkness-induced contractions (which are known to be mediated by cGMP[15]) but did not prevent NO-induced contractions, further supporting the existence of a second pathway. In other organisms, cGMP-independent NO signaling can involve S-nitrosation, the modification of proteins through the formation of an S-NO covalent bond,[51,52] although the direct targets and functions of S-nitrosation in animals are less well understood than NO/cGMP signaling. It is possible that this mechanism explains the cGMP-independent pathway underlying NO-induced colony contraction in *C. flexa*.

The control of multicellular behavior by NO/cCMP signaling in *C. flexa* is reminiscent of its function in animals, most notably in sponges.[53] In the demosponges *Tethya wilhelma*[1], *Ephydatia muelleri*[2], and *Spongilla lacustris*[3], NO induces global contractions and stops flagellar beating in choanocyte chambers, which interrupts feeding, allows expulsion of clumps of waste, and flushes the aquiferous canal system (a behavior sometimes called ‘‘sneezing’’).[54] Recently, single-cell RNA sequencing in *Spongilla lacustris* revealed that pinacocytes (epithelial cells that cover and shape the sponge body) co-express NOS and sGC,[3,55] the actomyosin contractility module and the transcription factor serum response factor (Srf), a master regulator of contractility.[3,56,57]

Control of motor and feeding behavior by NO/cGMP is also observed in cnidarians and some bilaterians. In the jellyfish *Aglantha digitale*, NO/cGMP signaling in neurons induces a switch from slow swimming (associated with feeding) to fast swimming (associated with escape) and inhibits tentacular ciliary beating.[10] In the sea pansy (a type of colonial cnidarian) *Renilla koellikeri*, NO/cGMP increases the amplitude of peristaltic contractions associated with the movement of body fluids through the gastrovascular cavity.[17] Finally, in the nudibranch *Clione limacine* and the snail *Lymnaea stagnalis*, NO activates both feeding and locomotory neural

circuits.[58–61] Thus, as in *C. flexa*, the ancient functions of NO/ cGMP signaling in animals may include the regulation of feeding and contraction.[16,22,61–64] Interestingly, NO signaling also controls metamorphosis in sponges,[9,65] gastropods,[66] annelids,[67] echinoderms,[68,69] and ascidians,[70,71] thus regulating a switch from swimming to feeding during irreversible developmental programs. The conservation of the NO-sensitive transduction pathway across choanozoans (including sGC, PKG, CNG, and PDEG; Figure A.2C) is consistent with a possible homology of the behavioral response to NO between choanoflagellates and animals. However, the mosaic distribution of eukaryotic and bacterial NOS across choanozoans (Figures 2.2A and A.1A) suggests that the source of NO itself might have switched an unknown number of times during evolution between the ancestral eukaryotic NOS, the horizontally transferred bacterial NOS, and exogenous sources.

In future, identifying the function of *Cf* NOS and *C. flexa* NO- or O_2-sensitive sGCs will require gene knockout, which is not yet possible in *C. flexa*. Moreover, studies on *Trichoplax* (in which NO/cGMP signaling has been predicted to exist based on genomic data[49]), additional animal phyla, and other choanoflagellates will help flesh out reconstitutions of the early evolution of animal NO signaling.

References

(1) Ellwanger, K.; Nickel, M. Neuroactive Substances Specifically Modulate Rhythmic Body Contractions in the Nerveless Metazoon *Tethya wilhelma* (Demospongiae, Porifera). *Front. Zool.* **2006**, *3* (1), 7

(2) Elliott, G. R. D.; Leys, S. P. Evidence for Glutamate, GABA and NO in Coordinating Behaviour in the Sponge, *Ephydatia muelleri* (Demospongiae, Spongillidae). *J. Exp. Biol.* **2010**, *213* (13), 2310–2321

(3) Musser, J. M.; Schippers, K. J.; Nickel, M.; Mizzon, G.; Kohn, A. B.; Pape, C.; Ronchi, P.; Papadopoulos, N.; Tarashansky, A. J.; Hammel, J. U.; Wolf, F.; Liang, C.; Hernández-Plaza, A.; Cantalapiedra, C. P.; Achim, K.; Schieber, N. L.; Pan, L.; Ruperti, F.; Francis, W. R.; Vargas, S.; Kling, S.; Renkert, M.; Polikarpov, M.; Bourenkov, G.; Feuda, R.; Gaspar, I.; Burkhardt, P.; Wang, B.; Bork, P.; Beck, M.; Schneider, T. R.; Kreshuk, A.; Wörheide, G.; Huerta-Cepas, J.; Schwab, Y.; Moroz, L. L.; Arendt, D. Profiling Cellular Diversity in Sponges Informs Animal Cell Type and Nervous System Evolution. *Science* **2021**, *374* (6568), 717–723

(4) Bogdan, C. Nitric Oxide Synthase in Innate and Adaptive Immunity: An Update. *Trends Immunol.* **2015**, *36* (3), 161–178

(5) Hillyer, J. F.; Estévez-Lao, T. Y. Nitric Oxide Is an Essential Component of the Hemocyte-Mediated Mosquito Immune Response against Bacteria. *Dev. Comp. Immunol.* **2010**, *34* (2), 141–149

(6) Gibbs, S. M.; Becker, A.; Hardy, R. W.; Truman, J. W. Soluble Guanylate Cyclase Is Required during Development for Visual System Function in *Drosophila*. *The Journal of Neuroscience* **2001**, *21* (19), 7705–7714

(7) Estephane, D.; Anctil, M. Retinoic Acid and Nitric Oxide Promote Cell Proliferation and Differentially Induce Neuronal Differentiation *in Vitro* in the Cnidarian *Renilla koellikeri*. *Dev. Neurobiol.* **2010**, *70* (12), 842–852

(8) Tomankova, S.; Abaffy, P.; Sindelka, R. The Role of Nitric Oxide during Embryonic Epidermis Development of *Xenopus laevis*. *Biol. Open* **2017**, *6* (6), 862–871

(9) Ueda, N.; Richards, G. S.; Degnan, B. M.; Kranz, A.; Adamska, M.; Croll, R. P.; Degnan, S. M. An Ancient Role for Nitric Oxide in Regulating the Animal Pelagobenthic Life Cycle: Evidence from a Marine Sponge. *Sci. Rep.* **2016**, *6* (1), 37546

(10) Moroz, L. L.; Meech, R. W.; Sweedler, J. V.; Mackie, G. O. Nitric Oxide Regulates Swimming in the Jellyfish *Aglantha digitale*. *Journal of Comparative Neurology* **2004**, *471* (1), 26–36

(11) Pirtle, T. J.; Satterlie, R. A. Cyclic Guanosine Monophosphate Modulates Locomotor Acceleration Induced by Nitric Oxide but Not Serotonin in *Clione limacina* Central Pattern Generator Swim Interneurons. *Integrative Organismal Biology* **2021**, *3* (1)

(12) Carr, M.; Leadbeater, B. S. C. Re-Evaluating Loricate Choanoflagellate Phylogenetics: Molecular Evidence Points to the Paraphyly of Tectiform Species. *Protist* **2022**, *173* (6), 125924

(13) Denninger, J. W.; Marletta, M. A. Guanylate Cyclase and the ·NO/CGMP Signaling Pathway. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1999**, *1411* (2–3), 334–350

(14) Andreakis, N.; D'Aniello, S.; Albalat, R.; Patti, F. P.; Garcia-Fernandez, J.; Procaccini, G.; Sordino, P.; Palumbo, A. Evolution of the Nitric Oxide Synthase Family in Metazoans. *Mol. Biol. Evol.* **2011**, *28* (1), 163–179

(15) Brunet, T.; Larson, B. T.; Linden, T. A.; Vermeij, M. J. A.; McDonald, K.; King, N. Light-Regulated Collective Contractility in a Multicellular Choanoflagellate. *Science* **2019**, *366* (6463), 326–334

(16) Colasanti, M.; Venturini, G.; Merante, A.; Musci, G.; Lauro, G. M. Nitric Oxide Involvement in *Hydra vulgaris* Very Primitive Olfactory-Like System. *The Journal of Neuroscience* **1997**, *17* (1), 493–499

(17) Anctil, M.; Poulain, I.; Pelletier, C. Nitric Oxide Modulates Peristaltic Muscle Activity Associated with Fluid Circulation in the Sea Pansy *Renilla koellikeri*. *J. Exp. Biol.* **2005**, *208* (10), 2005–2017

(18) Richter, D. J.; Fozouni, P.; Eisen, M. B.; King, N. Gene Family Innovation, Conservation and Loss on the Animal Stem Lineage. *eLife* **2018**, *7*

(19) Torruella, G.; de Mendoza, A.; Grau-Bové, X.; Antó, M.; Chaplin, M. A.; del Campo, J.; Eme, L.; Pérez-Cordón, G.; Whipps, C. M.; Nichols, K. M.; Paley, R.; Roger, A. J.; Sitjà-Bobadilla, A.; Donachie, S.; Ruiz-Trillo, I. Phylogenomics Reveals Convergent Evolution of Lifestyles in Close Relatives of Animals and Fungi. *Curr. Biol.* **2015**, *25* (18), 2404–2410

(20) Letunic, I.; Bork, P. Interactive Tree Of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49* (W1), W293–W296

(21) Barr, I.; Guo, F. Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *Bio. Protoc.* **2015**, *5* (18)

(22) Colasanti, M.; Persichini, T.; Venturini, G. Nitric Oxide Pathway in Lower Metazoans. *Nitric Oxide* **2010**, *23* (2), 94–100

(23) Robertson, H. M. The Choanoflagellate *Monosiga brevicollis* Karyotype Revealed by the Genome Sequence: Telomere-Linked Helicase Genes Resemble Those of Some Fungi. *Chromosome Research* **2009**, *17* (7), 873–882

(24) Fairclough, S. R.; Chen, Z.; Kramer, E.; Zeng, Q.; Young, S.; Robertson, H. M.; Begovic, E.; Richter, D. J.; Russ, C.; Westbrook, M. J.; Manning, G.; Lang, B. F.; Haas, B.; Nusbaum, C.; King, N. Premetazoan Genome Evolution and the Regulation of Cell Differentiation in the Choanoflagellate *Salpingoeca rosetta*. *Genome. Biol.* **2013**, *14* (2), R15

(25) Carr, M.; Leadbeater, B. S. C.; Hassan, R.; Nelson, M.; Baldauf, S. L. Molecular Phylogeny of Choanoflagellates, the Sister Group to Metazoa. *Proc. Natl. Acad. Sci. U S A* **2008**, *105* (43), 16641–16646

(26) Carr, M.; Richter, D. J.; Fozouni, P.; Smith, T. J.; Jeuck, A.; Leadbeater, B. S. C.; Nitsche, F. A Six-Gene Phylogeny Provides New Insights into Choanoflagellate Evolution. *Mol. Phylogenet. Evol.* **2017**, *107*, 166–178

(27) Derbyshire, E. R.; Marletta, M. A. Biochemistry of Soluble Guanylate Cyclase. In *Handbook of Experimental Pharmacology*; 2009; Vol. 191, pp 17–31

(28) Horst, B. G.; Stewart, E. M.; Nazarian, A. A.; Marletta, M. A. Characterization of a Carbon Monoxide-Activated Soluble Guanylate Cyclase from *Chlamydomonas reinhardtii*. *Biochemistry* **2019**, *58* (17), 2250–2259

(29) Gray, J. M.; Karow, D. S.; Lu, H.; Chang, A. J.; Chang, J. S.; Ellis, R. E.; Marletta, M. A.; Bargmann, C. I. Oxygen Sensation and Social Feeding Mediated by a *C. elegans* Guanylate Cyclase Homologue. *Nature* **2004**, *430* (6997), 317–322

(30) Huang, S. H.; Rio, D. C.; Marletta, M. A. Ligand Binding and Inhibition of an Oxygen-Sensitive Soluble Guanylate Cyclase, Gyc-88E, from *Drosophila*. *Biochemistry* **2007**, *46* (51), 15115–15122

(31) Cheung, B. H. H.; Arellano-Carbajal, F.; Rybicki, I.; de Bono, M. Soluble Guanylate Cyclases Act in Neurons Exposed to the Body Fluid to Promote *C. elegans* Aggregation Behavior. *Curr. Biol.* **2004**, *14* (12), 1105–1111

(32) Boon, E. M.; Marletta, M. A. Ligand Discrimination in Soluble Guanylate Cyclase and the H-NOX Family of Heme Sensor Proteins. *Curr. Opin. Chem. Biol.* **2005**, *9* (5), 441–446

(33) Boon, E. M.; Huang, S. H.; Marletta, M. A. A Molecular Basis for NO Selectivity in Soluble Guanylate Cyclase. *Nat. Chem. Biol.* **2005**, *1* (1), 53–59

(34) Maragos, C. M.; Morley, D.; Wink, D. A.; Dunams, T. M.; Saavedra, J. E.; Hoffman, A.; Bove, A. A.; Isaac, L.; Hrabie, J. A.; Keefer, L. K. Complexes of · NO with Nucleophiles as Agents for the Controlled Biological Release of Nitric Oxide. Vasorelaxant Effects. *J. Med. Chem.* **1991**, *34* (11), 3242–3247

(35) Zhao, Y.; Brandish, P. E.; DiValentin, M.; Schelvis, J. P. M.; Babcock, G. T.; Marletta, M. A. Inhibition of Soluble Guanylate Cyclase by ODQ. *Biochemistry* **2000**, *39* (35), 10848–10854

(36) Horst, B. G.; Marletta, M. A. Physiological Activation and Deactivation of Soluble Guanylate Cyclase. *Nitric Oxide* **2018**, *77*, 65–74

(37) Horst, B. G.; Yokom, A. L.; Rosenberg, D. J.; Morris, K. L.; Hammel, M.; Hurley, J. H.; Marletta, M. A. Allosteric Activation of the Nitric Oxide Receptor Soluble Guanylate Cyclase Mapped by Cryo-Electron Microscopy. *eLife* **2019**, *8*

(38) Doyle, M. P.; Duling, B. R. Acetylcholine Induces Conducted Vasodilation by Nitric Oxide-Dependent and -Independent Mechanisms. *American Journal of Physiology-Heart and Circulatory Physiology* **1997**, *272* (3), 1364–1371

(39) Sriram, K.; Laughlin, J. G.; Rangamani, P.; Tartakovsky, D. M. Shear-Induced Nitric Oxide Production by Endothelial Cells. *Biophys. J.* **2016**, *111* (1), 208–221

(40) Zhang, L.; Liu, Q.; Yuan, X.; Wang, T.; Luo, S.; Lei, H.; Xia, Y. Requirement of Heat Shock Protein 70 for Inducible Nitric Oxide Synthase Induction. *Cell Signal* **2013**, *25* (5), 1310–1317

(41) Rai, K. K.; Pandey, N.; Rai, S. P. Salicylic Acid and Nitric Oxide Signaling in Plant Heat Stress. *Physiol. Plant* **2020**, *168* (2), 241–255

(42) Giovine, M.; Pozzolini, M.; Favre, A.; Bavestrello, G.; Cerrano, C.; Ottaviani, F.; Chiarantini, L.; Cerasi, A.; Cangiotti, M.; Zocchi, E.; Scarfì, S.; Sarà, M.; Benatti, U. Heat Stress-Activated, Calcium-Dependent Nitric Oxide Synthase in Sponges. *Nitric Oxide* **2001**, *5* (5), 427–431

(43) Dulce, R. A.; Mayo, V.; Rangel, E. B.; Balkan, W.; Hare, J. M. Interaction Between Neuronal Nitric Oxide Synthase Signaling and Temperature Influences Sarcoplasmic Reticulum Calcium Leak. *Circ. Res.* **2015**, *116* (1), 46–55

(44) James-Clark, H. XXII.— *On the* Spongiæ Ciliatæ *as* Infusoria Flagellata; *or Observations on the Structure, Animality, and Relationship of* Leucosolenia Botryoides, *Bowerbank*. *Annals and Magazine of Natural History* **1868**, *1* (2), 133–142

(45) Andrews, G. *The Living Substance as Such: And as Organism*; 1897

(46) Nguyen, N. M.; Merle, T.; Broders, F.; Brunet, A.-C.; Sarron, F.; Jha, A.; Genisson, J.-L.; Rottinger, E.; Farge, E. Evolutionary Emergence of First Animal Organisms Triggered by Environmental Mechano-Biochemical Marine Stimulation. *bioRxiv*. Cold Spring Harbor Laboratory December 3, 2020, p 2020.12.03.407668

(47) C., L. B. S. Distribution and Chemistry of Microfilaments in Choanoflagellates, with Special Reference to the Collar and Other Tentacle Systems. *Protistologica* **1983**, *19*, 157–166

(48) Zhang, X. cGMP Signaling in Vertebrate Retinal Photoreceptor Cells. *Front. Biosci.* **2005**, *10* (1–3), 1191

(49) Moroz, L. L.; Romanova, D. Y.; Nikitin, M. A.; Sohn, D.; Kohn, A. B.; Neveu, E.; Varoqueaux, F.; Fasshauer, D. The Diversification and Lineage-Specific Expansion of

Nitric Oxide Signaling in Placozoa: Insights in the Evolution of Gaseous Transmission. *Sci. Rep.* **2020**, *10* (1), 13020

(50) Leadbeater, B. S. C. Life-History and Ultrastructure of a New Marine Species of *Proterospongia* (Choanoflagellida). *Journal of the Marine Biological Association of the United Kingdom* **1983**, *63* (1), 135–160

(51) Broniowska, K. A.; Diers, A. R.; Hogg, N. S-Nitrosoglutathione. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, *1830* (5), 3173–3181

(52) Smith, B. C.; Marletta, M. A. Mechanisms of S-Nitrosothiol Formation and Selectivity in Nitric Oxide Signaling. *Curr. Opin. Chem. Biol.* **2012**, *16* (5–6), 498–506

(53) Simion, P.; Philippe, H.; Baurain, D.; Jager, M.; Richter, D. J.; Di Franco, A.; Roure, B.; Satoh, N.; Quéinnec, É.; Ereskovsky, A.; Lapébie, P.; Corre, E.; Delsuc, F.; King, N.; Wörheide, G.; Manuel, M. A Large and Consistent Phylogenomic Dataset Supports Sponges as the Sister Group to All Other Animals. *Curr. Biol.* **2017**, *27* (7), 958–967

(54) Elliott, G. R. D.; Leys, S. P. Coordinated Contractions Effectively Expel Water from the Aquiferous System of a Freshwater Sponge. *J. Exp. Biol.* **2007**, *210* (21), 3736–3748

(55) Nickel, M.; Scheer, C.; Hammel, J. U.; Herzen, J.; Beckmann, F. The Contractile Sponge Epithelium *Sensu Lato* – Body Contraction of the Demosponge *Tethya wilhelma* Is Mediated by the Pinacoderm. *J. Exp. Biol.* **2011**, *214* (10), 1692–1698

(56) Miano, J. M.; Long, X.; Fujiwara, K. Serum Response Factor: Master Regulator of the Actin Cytoskeleton and Contractile Apparatus. *American Journal of Physiology-Cell Physiology* **2007**, *292* (1), C70–C81

(57) Brunet, T.; Fischer, A. H.; Steinmetz, P. R.; Lauri, A.; Bertucci, P.; Arendt, D. The Evolutionary Origin of Bilaterian Smooth and Striated Myocytes. *eLife* **2016**, 5:e19607

(58) Moroz, L. L.; Park, J.-H.; Winlow, W. Nitric Oxide Activates Buccal Motor Patterns in *Lymnaea stagnalis*. *Neuroreport* **1993**, *4* (6), 643–646

(59) Kobayashi, S.; Sadamoto, H.; Ogawa, H.; Kitamura, Y.; Oka, K.; Tanishita, K.; Ito, E. Nitric Oxide Generation around Buccal Ganglia Accompanying Feeding Behavior in the Pond Snail, *Lymnaea stagnalis*. *Neurosci. Res.* **2000**, *38* (1), 27–34

(60) Moroz, L. L.; Norekian, T. P.; Pirtle, T. J.; Robertson, K. J.; Satterlie, R. A. Distribution of NADPH-Diaphorase Reactivity and Effects of Nitric Oxide on Feeding and Locomotory Circuitry in the Pteropod Mollusc, *Clione limacina*. *J. Comp. Neurol.* **2000**, *427* (2), 274–284

(61) Moroz, L. L. Parallel Evolution of Nitric Oxide Signaling: Diversity of Synthesis and Memory Pathways. *Front. Biosci.* **2011**, *16* (1), 2008

(62) Cristino, L.; Guglielmotti, V.; Cotugno, A.; Musio, C.; Santillo, S. Nitric Oxide Signaling Pathways at Neural Level in Invertebrates: Functional Implications in Cnidarians. *Brain Res.* **2008**, *1225*, 17–25

(63) Yabumoto, T.; Takanashi, F.; Kirino, Y.; Watanabe, S. Nitric Oxide Is Involved in Appetitive but Not Aversive Olfactory Learning in the Land Mollusk *Limax valentianus*. *Learning & Memory* **2008**, *15* (4), 229–232

(64) Jacklet, J. W. Nitric Oxide Signaling in Invertebrates. *Invertebrate Neuroscience* **1997**, *3* (1), 1–14

(65) Song, H.; Hewitt, O. H.; Degnan, S. M. Arginine Biosynthesis by a Bacterial Symbiont Enables Nitric Oxide Production and Facilitates Larval Settlement in the Marine-Sponge Host. *Curr. Biol.* **2021**, *31* (2), 433-437.e3

(66) Froggett, S. J.; Leise, E. M. Metamorphosis in the Marine Snail *Ilyanassa obsoleta*, Yes or NO? *Biol. Bull.* **1999**, *196* (1), 57–62

(67) Biggers, W. J.; Pires, A.; Pechenik, J. A.; Johns, E.; Patel, P.; Polson, T.; Polson, J. Inhibitors of Nitric Oxide Synthase Induce Larval Settlement and Metamorphosis of the Polychaete Annelid *Capitella teleta*. *Invertebr. Reprod. Dev.* **2012**, *56* (1), 1–13

(68) Bishop, C. D.; Brandhorst, B. P. Development of Nitric Oxide Synthase-defined Neurons in the Sea Urchin Larval Ciliary Band and Evidence for a Chemosensory Function during Metamorphosis. *Developmental Dynamics* **2007**, *236* (6), 1535–1546

(69) Bishop, C. D.; Brandhorst, B. P. On Nitric Oxide Signaling, Metamorphosis, and the Evolution of Biphasic Life Cycles. *Evol. Dev.* **2003**, *5* (5), 542–550

(70) Ueda, N.; Degnan, S. M. Nitric Oxide Acts as a Positive Regulator to Induce Metamorphosis of the Ascidian *Herdmania momus*. *PLoS One* **2013**, *8* (9), e72797

(71) Comes, S.; Locascio, A.; Silvestre, F.; d'Ischia, M.; Russo, G. L.; Tosti, E.; Branno, M.; Palumbo, A. Regulatory Roles of Nitric Oxide during Larval Development and Metamorphosis in *Ciona intestinalis*. *Dev. Biol.* **2007**, *306* (2), 772–784

Chapter 3
Characterization of a NO-activated Homodimeric Soluble Guanylate Cyclase from *Choanoeca flexa*

Summary

Soluble guanylate cyclases are dimeric gas sensing enzymes in eukaryotes that catalyze the formation of cyclic GMP from GTP. While commonly studied sGCs from insects and vertebrates are heterodimers, there are additional classes of homodimeric sGCs in eukaryotes that are less well characterized. Herein we report the characterization of *Cf* sGC1 isolated from the organism *Choanoeca flexa*, a single-celled eukaryote. *Cf* sGC1 is a homodimeric sGC that exhibits a three-state activation profile in response to NO similar to that observed with heterodimeric NO-responsive sGCs. *Cf* sGC1 was isolated as an active homodimer, has one heme cofactor per dimer and exhibits typical saturation kinetics. Small angle X-ray scattering revealed that *Cf* sGC1 undergoes a structural change mirroring that of heterodimeric sGCs in the presence of excess NO (excess relative to the heme concentration). Additionally, the C-terminal catalytic domain of *Cf* sGC1 (sGC1-CAT) was expressed and characterized. The K_D for sGC1-CAT dimerization was 1.8 ± 0.4 μM. This highlighted the effect on enzyme dimerization by N-terminal domains.

This work was carried out in collaboration with Dr. William C. Thomas and Zachary B. Green in the lab of Prof. Michael A. Marletta. Y. W. and Z. B. G. performed biochemical characterization of *Cf* sGC1; Y. W. and W. C. T. performed SAXS experiments; W. C. T. performed data analysis for SAXS experiments.

Introduction

Nitric oxide (NO) is a crucial signaling molecule in mammals that mediates vasodilation, short-term memory formation, and platelet aggregation.[1–3] The cellular sensor for NO is the enzyme soluble guanylate cyclase (sGC). NO activation of sGC leads to the formation of the secondary messenger cyclic-3',5'-GMP (cGMP) via the cyclization of GTP.[4,5] Many sGCs from mammals and insects (α/β-type sGC) are heterodimeric and are selectively activated by NO through ligation at the heme prosthetic group. However, genomic and transcriptomic data analysis led to the discovery of homodimeric sGCs including some that are regulated by oxygen. [6–9] Homodimeric sGCs are similarly important for cell signaling, with examples like oxygen sensing sGCs from *Drosophila melanogaster* that modulate opening and closing of spiracles for gas exchange, and several sGCs expressed in *Caenorhabditis elegans* neurons have been found to mediate aerotaxis. [6,9–11]

The diatomic gas-sGC-cGMP signaling pathway is present in diverse animal lineages. That said, sGC signaling has been less studied in non-mammalian species, leaving still under characterized aspects of NO signaling that may hold important insight into how it evolved, as well as how sGC activity is regulated through ligand-induced conformational change. The first non-animal sGC to be biochemically characterized, *Cr* Cyg-11, was from the algae *Chlamydonomas reinhardtii*. It is distinct from α/β-type sGCs in that it is activated by carbon monoxide to a greater extent than by NO.[10] The O_2-binding, homodimeric sGC Gyc-88E from *Drosophila melanogaster*,

on the other hand, is inhibited by O_2 binding.[6] sGC Gyc-88E is not selective for NO and it is also inhibited by NO. The examples above highlight the diversity of activity response to ligand binding that subclasses of sGC proteins can exhibit and suggest distinctive but related underlying mechanisms of activity regulation through ligand-induced conformational change.

Recently we described a sGC-mediated, NO-sensitive collective contraction phenotype in the choanoflagellate species *Choanoeca flexa*. Choanoflagellates are the closest unicellular relatives of animals, and they utilize diverse signaling pathways that are conserved across choanoflagellates and animals to regulate the morphology and activity of single cells and multicellular colonies.[12,13] As a result, several choanoflagellate species have been developed as model organisms to study the molecular evolution of animal multicellularity.[14,15] In *C. flexa*, nutrient depletion studies and treatment with phosphodiesterase inhibitors have shown that a rhodopsin-controlled cGMP phosphodiesterase signals the transformation between two distinct colony morphologies.[16] It was further demonstrated that this process can be triggered by the addition of NO and inhibited in the presence of the sGC inhibitor ODQ.[17,18] Taken together, these findings suggest that a NO-activated sGC is responsible for signaling colony transformation and prompted a search for sGC proteins encoded in the *C. flexa* genome.

Based on existing transcriptomic data, four genes encoding sGCs (named *Cf* sGC1 through 4) have been identified in *C. flexa*, with each containing a predicted heme nitric oxide oxygen-binding (H-NOX) domain, a Per-Arnt-Sim (PAS) like domain, a coiled-coiled region, and a catalytic cyclase domain (CAT), mirroring the canonical domain architecture of previously characterized sGCs (Figure 3.1A). Alignment of the H-NOX domain of these four sGC sequences revealed that three contained a tyrosine residue in the distal ligand binding pocket of the H-NOX domain that is associated with the ability to bind O_2, and, therefore, could be regulated by O_2.[6,19] *Cf* sGC1, the only sGC identified from *C. flexa* that does not contain a H-NOX domain distal tyrosine residue, was predicted to be selective for NO regulation and is the focus of studies reported here (Figure 3.1B).

Our previous work reported a preliminary characterization of *Cf* sGC1.[17] We found that *Cf* sGC1 is capable of binding NO and CO but does not bind O_2. We also found that *Cf* sGC1 is active as a homodimer and exhibits a 3-stage activation process reminiscent of a α/β-type sGC, making *Cf* sGC1 the first homodimeric sGC reported with a α/β-type sGC-like activity profile. Herein a detailed biochemical characterization of *Cf* sGC1 is presented. Substrate kinetics was measured for *Cf* sGC1 under different ligand bound states, the oligomeric state was determined to be a homodimer, and the heme stoichiometry was determined to be on average one heme per homodimer. Small angle X-ray scattering was employed to demonstrate that *Cf* sGC1 is activated through a conformational change that involves extension of the protein structure, a feature previously only reported in α/β-type sGC.

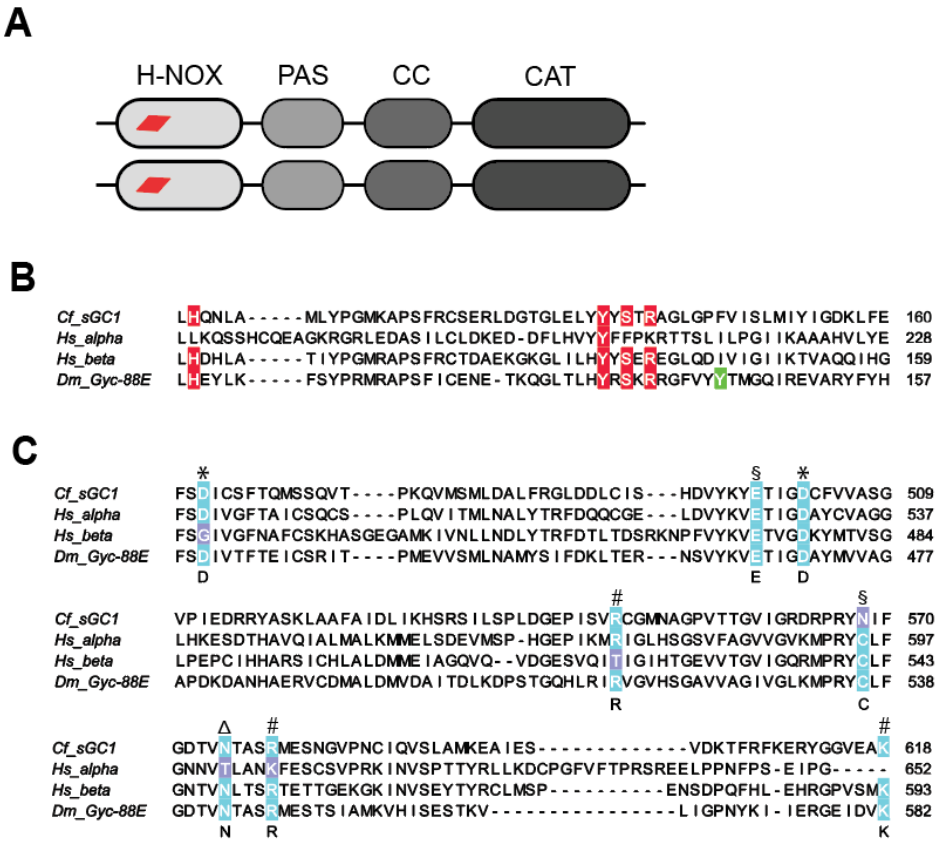


Figure 3.1. Domain arrangement and multiple sequence alignment of *Cf* sGC1**.** (A) Illustration of the predicted domain architecture of *Cf* sGC1. *Cf* sGC1 shares domain architecture with canonical sGCs, containing the Heme-Nitric oxide/OXygen binding (H-NOX), Per-Arnt-Sim (PAS), coiled coil (CC), and catalytic (CAT) domains. The predicted heme binding sites are indicated with a red parallelogram. (B) Multisequence alignment of *Cf* sGC1 to human heterodimeric sGC (*Hs*_alpha and beta) and the O_2 binding homodimeric sGC from the fruit fly *Drosophila melanogaster*, Gyc-88E. Highlighted in red are the His residue in the proximal pocket and the YxSxR motif for interacting with the heme propionate side chains. Highlighted in green is the tyrosine residue for stabilizing the heme Fe(II)-O_2 complex. (C) Sequence alignment of the CAT domain of *Cf* sGC1 to human sGC and fly sGC. Highlighted are required residues for a complete catalytic active site[20] and the types of residues required for function are labelled underneath each line. The symbols used are: metal binding (*), guanine binding (§), phosphate binding (#) and ribose binding (Δ). Alignment was prepared using Clustal Omega Multisequence Alignment Tool and visualized using JalView 2.11.

Materials and methods

Materials

Primers were purchased from Integrated DNA Technologies. Gibson Master Mix and *E. coli* competent cell stocks were purchased from the UC Berkeley QB3 MacroLab. PrimeStar MAX DNA polymerase master mix was purchased from Takara Bio USA. Plasmid extraction and DNA purification kits were purchased from Qiagen. Terrific Broth media powder, kanamycin, benzamidine and 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) were purchased from Research Products International. Sodium dithionite ($Na_2S_2O_4$), DNase I from bovine pancreas, sodium phosphate, sodium acetate, magnesium acetate, carbon monoxide (CO) gas and guanosine 5'-triphosphate sodium salt were purchased from Sigma Aldrich. Isopropyl β-

D-1-thiogalactopyranoside (IPTG), 4-(2-hydroxyethyl)piperazine ethanesulfonic acid (HEPES), Pierce BCA Protein Assay Kit, ammonium acetate and acetonitrile (HPLC grade) were purchased from Thermo Fisher Scientific. Lysozyme from chicken egg white and sodium chloride were purchased from VWR Scientific. Dithiothreitol (DTT) was purchased from BACHEM. Imidazole was purchased from Oakwood Chemical. Vivaspin spin concentrators were purchased from Sartorius. D-Aminolevulinic acid (d-ALA) and hemin chloride were purchased from Spectrum Chemical Company. proliNONOate was purchased from Cayman Chemical Company.

Protein expression and purification

pET_*Cf* sGC1 and pET_*Cf* sGC1_CAT were transformed into *E. coli* BL21Star (DE3) competent cells. Overnight culture was prepared by inoculating one colony in 50 mL of Luria Broth media supplemented with 50 µg/mL kanamycin. After overnight shaking at 37 °C, the overnight culture was diluted 1:200 into Terrific Broth (TB) media supplemented with 50 µg/mL kanamycin and 2% glucose. Additionally, for expressing *Cf* sGC1, the overnight culture and the large culture were supplemented with 500 µM $FeCl_3$. The large culture in TB media was incubated at 37 °C until cell density reached OD_{600} = 0.6 – 0.8 before protein production was induced by the addition of 500 µM IPTG. For expressing full length *Cf* sGC1, d-ALA (1 mM) was supplemented when the cell density of the culture reached OD_{600} = 0.4 – 0.6 and the culture was incubated at 37 °C for an additional 15 minutes before protein production was induced. After induction, the culture was incubated at 18 °C for 16 – 20 hours, and the cells were harvested through centrifugation at 4500 x g, 4 °C for 25 min. The cell paste was collected, flash frozen in liquid nitrogen, and transferred to -80 °C for storage.

Protein purification was carried out at 4 °C unless otherwise specified. Cell pellets were thawed in an ice water bath and resuspended in ice cold Buffer A (50 mM sodium phosphate, 150 mM NaCl, 5 mM imidazole, 5 % glycerol, pH 8.0) supplemented with 50 mM benzamidine, 0.2 mM AEBSF, 0.25 mg/mL DNAse I, and 0.25 mg/mL lysozyme. The cell resuspension was lysed using a high-pressure homogenizer (Avestin Emulsiflex C5). The resulting cell lysate was clarified by centrifugation (36,000 x g, 55 min) and passed through a gravity column loaded with His60 Ni Superflow Resin equilibrated with Buffer A. The column was subsequently washed with 10 CV Buffer A, 10 CV Buffer B (50 mM sodium phosphate, 200 mM NaCl, 40 mM imidazole, 5 % glycerol, pH 8.0) and His-tagged protein was eluted with 5 CV Buffer C (50 mM sodium phosphate, 200 mM NaCl, 400 mM imidazole, 5 % glycerol, pH 8.0). Protein-containing fractions determined by SDS-PAGE electrophoresis were combined and concentrated using a 50 kDa molecular weight cut-off spin concentrator and diluted 10x in Buffer D (25 mM triethanolamine, 25 mM NaCl, 5 mM DTT, pH 7.4). The diluted protein was loaded on to a POROS HQ20 anion exchange column (Thermo Scientific) equilibrated with Buffer D and the protein was eluted over a gradient of 5% to 50% Buffer E (25 mM triethanolamine, 750 mM NaCl, 5 mM DTT, pH 7.4) over 17 CV at 5 mL/min. Fractions were separated into main and side fractions where fractions with a ratio of A_{428}:A_{280} > 1 were considered main fractions and combined, concentrated to 25 – 50 µM and flash frozen in liquid nitrogen. Purified protein was verified using mass spectrometry (see appendix B). Protein aliquots were stored at -80 °C for future use.

Small-angle X-ray scattering

Small-angle X-ray scattering (SAXS) was performed on *Cf* sGC1 under anaerobic conditions in a Coy glovebox at the MacChess BioSAXS beamline (Sector 7A) at the Cornell

High-Energy Synchrotron Source, Ithaca, New York.[21–23] For both samples, *Cf* sGC1 was thawed and reduced directly in the glovebox, then exchanged into SAXS assay buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1% glycerol, 1 mM TCEP) using a pre-equilibrated Zeba spin desalting column (Thermo Scientific). Heme reduction was confirmed via in-line UV-vis spectroscopy (AvaSpec-ULS2048, Avantes), with the UV flow cell located directly after the SAXS sample cell. For the unliganded *Cf* sGC1 sample, 100 µL of ~100 µM sample was injected onto a Superdex 200 10/300 column (Cytiva Life Sciences) pre-equilibrated with assay buffer, and the column was run using an external pump at 0.5 mL/s (LC-20AD, Shimadzu). The xsNO sample was performed under similar conditions but with the addition of small molecule NO releasing agents, referred to as NONOates.[24] 100 µM proliNONOate was added directly to the sample prior to loading. Further, to ensure that the protein did not exchange into NO-free buffer on the column, the column was pre-equilibrated with A: 100 µM DETA-NONOate (half-life 56 hours at 22 – 25 °C) B: 100 µM DEA-NONOate (half-life 16 min at 22 – 25 °C) 40 minutes before loading.

Data were collected using a Dectris EIGER 4M detector at 4 °C. 1,500 × 2s X-ray exposures were collected for each run. Scattering images were integrated about the beam center and normalized by transmitted intensities measured on a photodiode beamstop. The X-ray wavelength of the experiment was λ = 1.1061 Å and the sample-to-detector distance was 1,748.0 mm, as determined by silver behenate calibration. Scattering was collected over a range of q = 0.01 Å^{-1} to 0.3 Å^{-1}, where q is the scattering vector $q = 4\pi\sin\theta/\lambda$ and 2θ is the scattering angle. Data processing, analysis, and comparison to predicted and published scattering profiles were performed in BioXTAS RAW using established protocol.[25] Background subtraction was performed by subtracting the buffer baseline prior to peak elution. Radii of gyration (R_g) and molecular weight were estimated on a per frame basis using Guinier analysis and the volume of correlation method, respectively, as implemented in RAW.[26,27] Principal components of the elution were deconvolved using Evolving Factor Analysis (EFA),[28] which enables mathematical separation of partially co-eluting species using iterative singular value decomposition to identify the number of distinguishable eluting species. An initial component eluting prior to the main peak was identified as minor sample aggregation. The main component was also computationally separated from a smaller species eluting in the shoulder of the main peak. Pair distance distribution analysis was performed using a Bayesian Indirect Fourier Transform (BIFT).[29] Error bars associated with R_g values are curve-fitting uncertainties from Guinier analysis. The calculated scattering of the predicted *Cf* sGC1 homodimer was determined using Crysol and validated with FoXS.[30–32] Low-resolution electron density maps were calculated using DENSS, as implemented in RAW.[33]

Predicted structure model

The sequence of *Cf* sGC1 was used as input using the Google AlphaFold web server powered by AlphaFold 3.[34] Default settings for multimeric protein predictions were used. The rank order 1 structure was used to calculate predicted SAXS scattering, but no significant differences were observed between different rank order predictions. To predict a potential asymmetric conformation of *Cf* sGC1, the Swiss-Model homology-model server was used.[35] The sequence of *Cf* sGC1 was used as input and aligned with the top hits for both β and α chains corresponding to ligand-free and NO-bound protein from the protein databank, which constitute the human sGC β and α subunits in the ligand-free (PDB:7D9R) and NO-bound (PDB:8HBH) states.[36] *Cf* sGC1 monomers models were then built in Swiss-Model, and monomers corresponding to human α and β were then combined to form dimer models.

Heme content quantification

Pyridine hemochromagen assay was carried out following a published protocol with minor modifications to quantify the heme content of *Cf* sGC1 protein.[37] First, *Cf* sGC1 protein was desalted into Buffer F (50 mM HEPES, 150 mM NaCl, 5 % glycerol, pH 7.5, 0.22 μm filtered) and diluted to 4 times the original volume. The diluted protein sample (100 μL) was then mixed with Solution 1 (100 μL, 0.2 M NaOH, 40 % pyridine, 5 mM potassium ferricyanide) to yield the oxidized pyridine hemochrome sample. An aliquot of solution 2 (2 μL, 0.5 M sodium dithionite, 0.5 M NaOH) was then added to yield the reduced pyridine hemochrome sample. The UV-vis absorption spectrum of the reduced pyridine hemochrome was measured on a Cary 300 UV-vis spectrophotometer (Agilent). The absorption spectrum was baseline corrected against a UV-vis absorption spectrum of a sample composed of Buffer F and Solution 1 mixed in a 1:1 ratio, and the baseline corrected absorbance at 557 nm (ε = 34,700 $M^{-1} \cdot cm^{-1}$) of the reduced pyridine hemochrome was used to calculate the heme concentration in the *Cf* sGC1 sample.

Activity assay and quantification

Initial rates of *Cf* sGC1 were measured by quantifying the production of cGMP in endpoint assays performed at 25 °C under anoxic conditions. All steps preparing *Cf* sGC1 prior to the activity assay were carried out at 4 °C unless otherwise noted. *Cf* sGC1 was reduced using 5 mM $Na_2S_2O_4$ and desalted into Ar-sparged Buffer F using a pre equilibrated 7.5 kDa MWCO Zeba Spin Desalting column (Thermo Scientific). Concentration of the protein was determined by the absorbance at 428 nm (ε = 101,000 M^{-1} cm^{-1}).[17] xsNO protein sample was prepared by the addition of 250 μM proliNONOate. The sample was incubated for 5 minutes, then desalted into Buffer F using a preequilibrated spin desalting column to yield the 1-NO protein sample. Activity assay mixtures were prepared in Ar-sparged Buffer F supplemented with 5 mM DTT and $MgCl_2$ greater or equal to two-fold the concentration of GTP and included 50 nM *Cf* sGC1 protein. For the xsNO assay, NO was supplemented by the addition of 50 μM proliNONOate. CO-saturated Buffer F was prepared by sparging Buffer F using CO gas for 15 minutes, and it was used to prepare the reaction mixture for the CO-bound assay sample. Reaction was initiated by the addition of GTP and aliquots of the reaction mixture were removed and quenched using 1 % formic acid. The results of the enzymatic reaction were analyzed using reverse phase HPLC as previously reported.[10] Briefly, the reaction timepoint samples were injected onto an Eclipse Plus C18 column (4.6 x 100 mm, 3.5 μm particle size, Agilent). A gradient composed of 20 mM ammonium acetate, 0.1 % formic acid (Buffer G) and 99.9 % acetonitrile, 0.1 % formic acid (Buffer H) was used to separate GTP and cGMP. The gradient was as follows: 0 – 6 min, 2% Buffer H; 6 – 7.5 min, 25 – 100% Buffer H; 7.5 – 8.5 min, 100 % Buffer H. Concentrations of cGMP were determined from the peak area eluting at 4.2 minutes using a standard curve constructed from 1.25 – 20 μM cGMP standards.

Initial rates of sGC1-CAT were also measured using similar endpoints assays at 25 °C as described above with minor modifications to the assay procedures. Assays for sGC1-CAT were carried out under aerobic conditions. Enzyme concentration was quantified using the absorbance at 280 nm using the calculated dimeric extinction coefficient of 21,610 $M^{-1} \cdot cm^{-1}$. In the endpoint assays for the determination of substrate kinetics parameters, the reaction mixture was prepared with 5 μM sGC1-CAT protein. In the endpoint assays for the determination of sGC1-CAT dimer affinity, the reaction mixture was prepared with 2 mM GTP and 5 mM $MgCl_2$.

Analytical size exclusion chromatography

Analytical size exclusion chromatography was carried out using a Superdex 200 Increase 10/300 GL column (GE Healthcare) pre-equilibrated with Buffer D. The elution profile of the sGC1-CAT protein (2.5 mg/mL) was compared to the elution profile of a set of protein standards dissolved in Buffer D at 0.2 mg/mL: Thyroglobulin (670 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66.5 kDa), DNase I (31 kDa) and lysozyme (14.3 kDa). The retention time was plotted against the log of the molecular weight, and a standard curve was prepared to determine the estimated molecular weight of sGC1-CAT.

Results

Oligomeric state and heme stoichiometry of Cf sGC1

The C-terminal hexa histidine-tagged *Cf* sGC1 was expressed in an *E. coli* BL21Star (DE3) heterologous host. Heme precursors d-ALA and iron (III) chloride were supplemented to facilitate heme overproduction. *Cf* sGC1 was purified to homogeneity using Ni-affinity chromatography followed by anion exchange chromatography. The molecular weight of the purified protein was verified using mass spectrometry (Figure B.1). To show that *Cf* sGC1 is a homodimer as predicted, SEC-SAXS was performed. The protein eluted as a single major peak with a small shoulder indicative of aggregation eluting prior to the protein and a longer shoulder eluting after the main peak (Figure 3.2A). The molecular weight as estimated by volume of correlation on a per frame basis was found to average ~165 kDa for the main peak, which best matches a dimeric species (expected MW 150 kDa). Evolving Factor Analysis was used to computationally separate the main peak (Figure B.2), and the resulting SAXS profile was found to be similar to that of a prototypical α/β-type sGC, *Manduca sexta* sGC.[25] The scattering profile of the main peak was calculated to have a $R_g = 48.2 \pm 0.2$ Å. When plotted as a Kratky plot, the scattering approaches zero at high q, indicating that *Cf* sGC1 obtained through heterologous expression in *E. coli* is a well folded, globular protein (Figure 3.3A).[22]

The shoulder region eluting after the main peak, here referred to as Component 2, was found to be at significantly lower concentration. It also was found to have a R_g smaller than that expected for an sGC dimer ($R_g = 38.5 \pm 0.6$ Å). The in-line UV-vis spectrum of the elution shoulder corresponding to Component 2 shows a marked decrease in absorbance at 426 nm when compared to the main peak (Figure B.2). This peak may represent truncated or apo-sGC monomers and is likely an artifact from heterologous expression and purification.

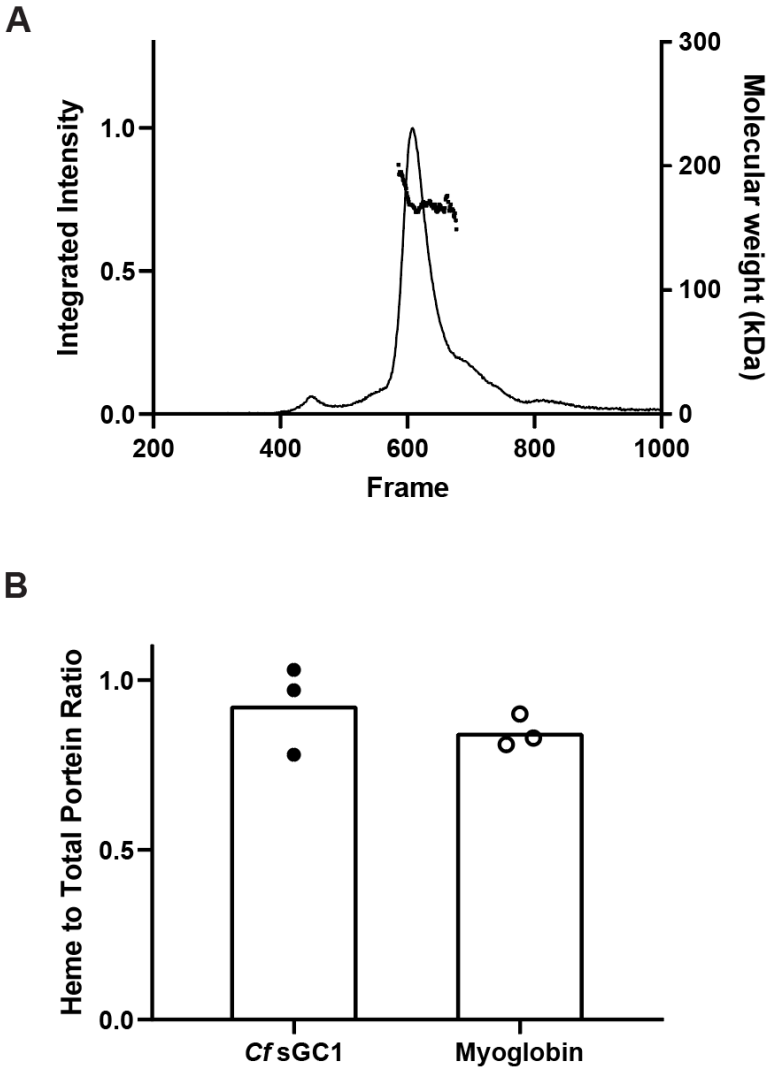


Figure 3.2. Oligomeric state and heme binding of *Cf* sGC1. (A) Size exclusion chromatography – coupled small angle X-ray scattering (SEC-SAXS) elution trace and estimated R_g of each scattering frame (dots) indicate that the *Cf* sGC1 is a homogeneous dimer of R_g = 48.2 ± 0.2 Å. (B) The results from pyridine hemochromagen assay coupled with bicinchoninic acid assay are consistent with on average one heme per *Cf* sGC1 homodimer. Myoglobin was used as a standard.

Homodimeric *Cf* sGC1 contains two H-NOX domains, so in principle, each *Cf* sGC1 homodimer could contain two heme cofactors. To quantify the heme stoichiometry, the pyridine hemochromagen assay was used to measure heme content in conjunction with bicinchoninic acid assay for protein determination. Contrary to prediction, the results were consistent with an average of one heme per homodimer (Figure 3.2B). Given that the prediction was two hemes per dimer, heme reconstitution was attempted by incubation of the protein with excess heme. The UV-vis spectrum after incubation and removal of the excess heme indicated the presence of non-specifically bound heme even after the second gel filtration step, and the subsequent activity assay showed a decrease in activity, suggesting that the non-specific bound heme negatively affected the normal activity of the protein (Figure B.5). Because of the spectral overlap between the non-specifically bound heme absorbance and the Soret band absorbance of the specifically bound heme, we were unable to determine if reconstitution increased specific heme binding. Reconstitution

assays were not pursued further. Taken together, *Cf* sGC1 contains on average one heme cofactor per homodimer.

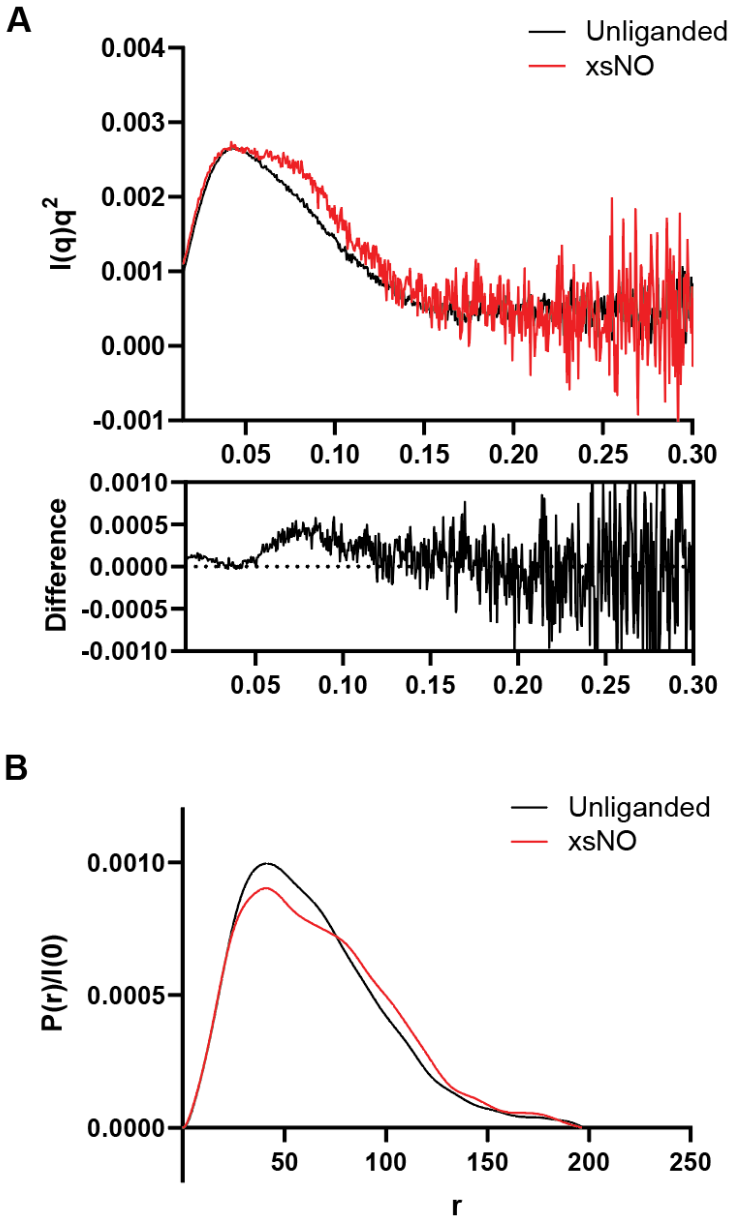


Figure 3.3. SAXS results of *Cf* sGC1 under unliganded (black) and xsNO (red) conditions. Results are consistent with a subtle conformational elongation when the protein is in the presence of excess NO. (A) SAXS profiles are shown in a Kratky plot to emphasize mid q features. When compared to the experimentally determined *Ms* sGC scattering profile, both species exhibit an increase in mid-q scattering with NO bound. The residual plot indicates a subtle shift near q = 0.1, indicating an extension of conformation when the protein encounters NO. (B) The electron pair distribution function indicates a slightly elongated pair distribution, suggesting conformational elongation.

Ligand bound activity and substrate kinetics of Cf sGC1

Like α/β-type sGC, *Cf* sGC1 forms a stable 5-coordinate high spin Fe(II) heme-NO complex (Figure B.3). Furthermore, *Cf* sGC1 has been shown to have a three stage activity profile in response to NO binding, where (*i*) without NO, the protein had a low, basal level activity; (*ii*)

with 1 molar equivalent of NO bound at the heme cofactor, the protein was activated ~2 fold, and (*iii*) in the presence of more than 2 molar equivalents of NO (xsNO), the protein was fully activated to ~5 fold its basal activity.[17] α/β-type sGC are also activated by carbon monoxide (CO) to 1-NO activity levels.[38] The effect of CO on the activity of *Cf* sGC1 was compared to the NO-bound activity states (Figure 3.4A). In the presence of more than equimolar CO with respect to heme, *Cf* sGC1 showed specific activity of 87.0 ± 6.7 nmol/min/mg, which was in a similar range as the 1-NO state activity of 66.6 ± 17.9 nmol/min/mg.

The steady state kinetics of *Cf* sGC1 was determined under unliganded, 1-NO and xsNO conditions. *Cf* sGC1 showed saturation kinetics under xsNO condition and showed a minor deviation from saturation at high, non-physiological GTP concentrations under the unliganded and 1-NO conditions (Figure 3.4B, Table 3.1). *Cf* sGC1 showed an increase in V_{max} with NO bound but did not show a significant change in the K_M. This is a noteworthy difference when compared to the α/β-type sGC, where the K_M for GTP is significantly decreased upon NO binding.

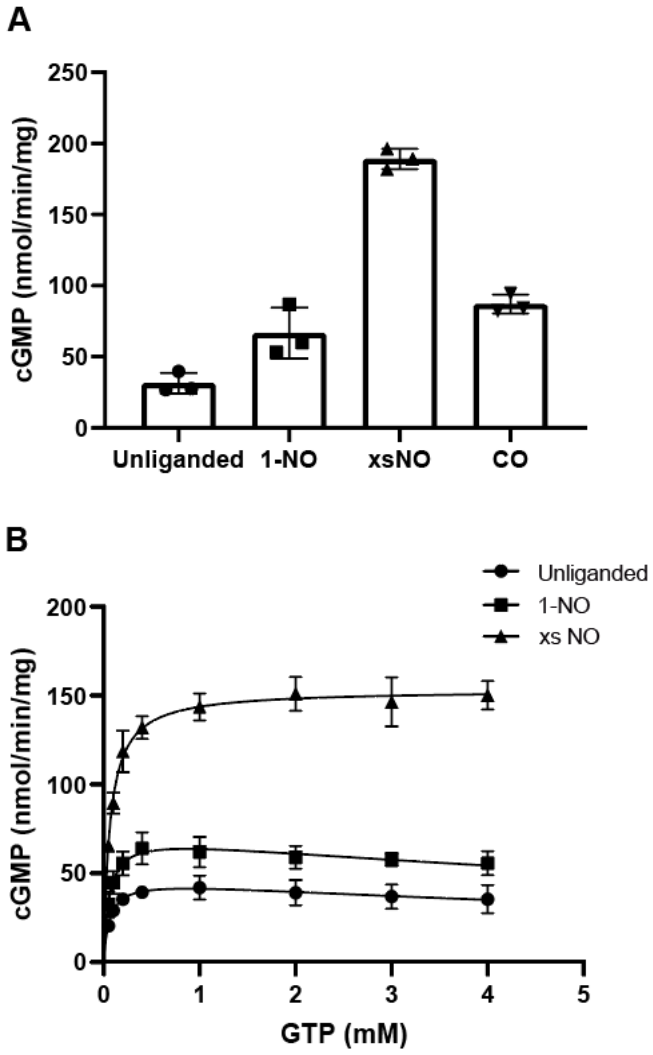


Figure 3.4. Activity and steady state kinetics of *Cf* sGC1. (A) Activity of *Cf* sGC1 under various ligation states. Activity at 1.5 mM GTP and 3 mM $MgCl_2$ was reported. Data shown are the averages of three replicates, and the error bars indicate one standard deviation. (B) Steady state kinetics of *Cf* sGC1. Initial rate of the reaction was measured at varying substrate concentrations and fitted to the saturation model or substrate inhibition model. Data shown is an average of four replicates, and error bars indicate one standard deviation.

Table 3.1: Steady state kinetics parameters of *Cf* sGC1 under basal, equimolar NO bound, and excess NO conditions. Deviation from saturation of U and 1-NO state protein occurs at high, non-physiological GTP concentrations.

	K_M (μM)	V_{max} (nmol/min/mg)	K_i (mM)	k_{cat} (s^{-1})
U	64.1 ± 15.6	47.4 ± 3.3	11.9 ± 4.9	0.12
1-NO	59.6 ± 12.9	72.5 ± 4.4	12.5 ± 4.6	0.18
xsNO	66.4 ± 5.5	153.2 ± 2.2		0.39

NO induces a conformational change in Cf sGC1

In previously studied heterodimeric sGCs, NO binding triggers a conformational change from an inactive, compact form to an active, extended form. To better understand potential conformational transitions required for *Cf* sGC1 activation, SEC-SAXS experiments on the protein in the presence of excess NO were carried out under anoxic conditions.

Like unliganded *Cf* sGC1, NO-bound *Cf* sGC1 elutes in a single primary peak with a small amount of aggregate and a post-elution shoulder (Figure B.3A). The principal component of the SEC-SAXS elution of NO-bound *Cf* sGC1 is slightly larger than unliganded *Cf* sGC1 (R_g = 50.7 ± 0.5 Å). When SAXS profiles are directly compared, a subtle shift in mid-*q* scattering is visible, denoting an opening of the structure (Figure 3.3A). This is further evidenced by a distinct shift in the pair distance distribution function observed in the region between 40 and 90 Å (Figure 3.3B). By comparison, the lower-concentration Component 2 species does not significantly change in R_g (NO-free: 38.5 ± 0.6 Å, compared to NO-present: 39.8 ± 1.0 Å) or SAXS profile shape upon addition of NO (Figure B.2).

The SAXS profiles of the main species of *Cf* sGC1 in both the NO-free and NO-bound experiments do not match the calculated scattering of an AlphaFold 3 predicted *Cf* sGC1 homodimer (Figure B.4). Instead, the SAXS profiles of NO-present *Cf* sGC1 more closely resembles that of NO-present *M. sexta* sGC, a prototypical heterodimer, where only a fraction of the protein in the SAXS sample was in an extended conformation.[25] The ligand-free *Cf* sGC1 SAXS profile more closely resembles the ligand-free *M. sexta* sGC profile, albeit with a more compact Kratky profile that may indicate a less extended inactive state than previously observed in α/β-type sGC. These similarities between *Cf* sGC1 and α/β-type sGCs are further supported by DENSS *ab initio* 3D reconstructions of *Cf* sGC1 density of the SAXS profiles of unliganded sGC1 and xsNO sGC1, which more closely resemble inactive and active α/β-type sGCs, respectively (Figure B.5).

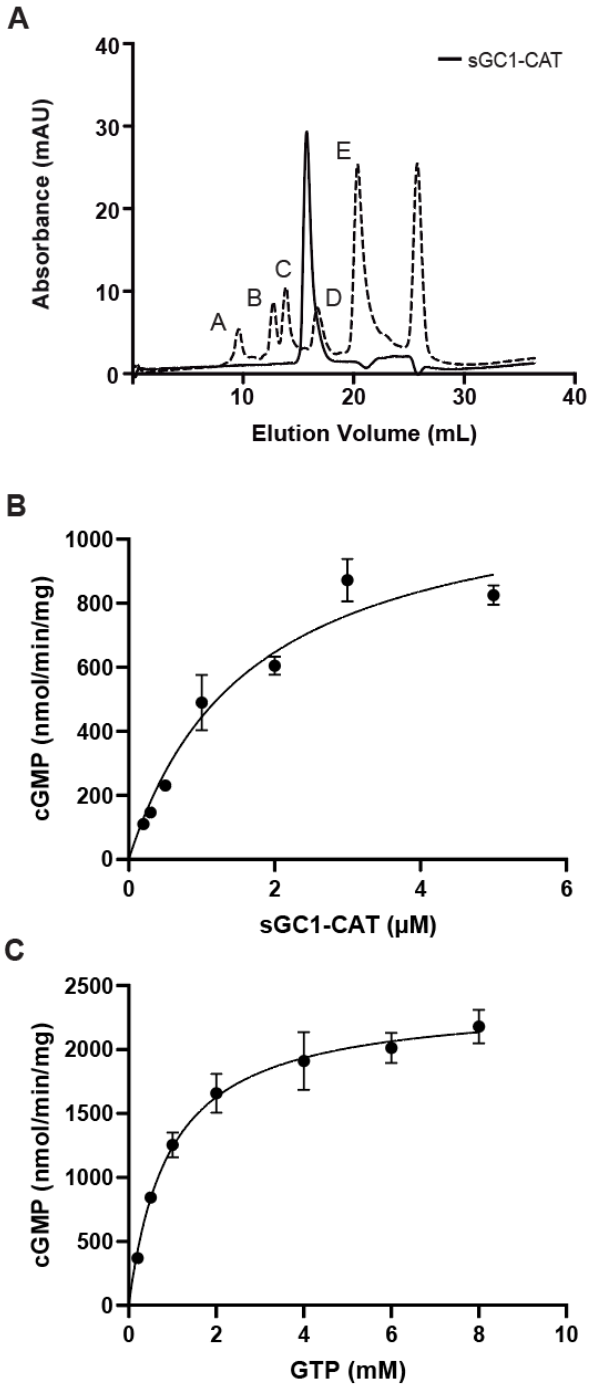


Figure 3.5. Dimer affinity and kinetics of sGC1-CAT dimers. (A) Analytical size exclusion chromatography of sGC1-CAT compared to elution profile of a set of five globular protein standards indicates that it eluted as a homodimer. Protein standards: A, thyroglobin, 670 kDa; B, alcohol dehydrogenase, 150 kDa; C, bovine serum albumin, 66.5 kDa; D, DNAse I, 31 kDa; E, lysozyme, 14 kDa. (B) Initial rates were measured for CAT domain at varying enzyme concentrations as a proxy for presence of active protein dimers. Curve was fitted to a standard dissociation function. $K_d = 1.8 \pm 0.4$ μM. (C) Kinetics of sGC1-CAT were collected by measuring initial rates at varying GTP concentrations. CAT domain exhibits Michaelis-Menten kinetics. $K_M = 0.93 \pm 0.10$ mM; $V_{max} = 2382 \pm 67$ nmol/min/mg; $k_{cat} = 2.2$ s^{-1}.

Properties of the catalytic domain of Cf sGC1

To gain further understanding into the activity regulation of *Cf* sGC1, a C-terminal catalytic domain truncation of *Cf* sGC1 (sGC1-CAT) was cloned, expressed and biochemically characterized. First, sGC1-CAT was purified to homogeneity through Ni-affinity chromatography, SEC, and anion exchange chromatography. Mass spectrometry was used to verify the molecular weight of the purified sGC1-CAT protein (Figure B.1B). Analytical SEC indicated that the sGC1-CAT eluted as a homodimer with an estimated molecular weight of 55.2 kDa, consistent with the expected molecular weight of 53.4 kDa (Figure 3.5A, Figure B.8). In addition to analytical SEC,

the dimer affinity of sGC1-CAT was determined by assaying the enzyme activity at increasing protein concentrations. Guanylate cyclases are only active as dimers, because a complete catalytic active site requires residues from both subunits. Therefore, the activity can be used as a proxy for determining the portion of dimeric enzyme in solution. The K_d of sGC1-CAT monomer to dimer formation was determined to be 1.8 ± 0.4 μM (Figure 3.5B). This value is on the same order of magnitude as the K_d of the CAT domain of α/β-type sGCs.[39] Finally, the steady state kinetics of sGC1-CAT homodimer was determined (Figure 3.5C). The sGC1-CAT homodimer displayed saturation kinetics with K_M = 0.93 ± 0.10 mM and V_{max} = 2382 ± 67 nmol/min/mg.

Discussion

This work presents the characterization *Cf* sGC1, the first homodimeric sGC characterized that has a α/β-type sGC-like activation profile. *Cf* sGC1 is also the first characterized choanoflagellate-derived sGC. Choanoflagellates are the closest unicellular living relatives of animals and are studied to understand the evolution of cellular pathways involved in animal multicellularity.[12] Our findings provide biochemical evidence of NO signaling in choanoflagellates and has further implications relevant to understanding NO signaling in early branching animal lineages, such as sponges and cnidarians. Our data is best viewed in comparison to previously well characterized α/β-type sGCs and Cyg-11 from the green algae *Chlamydomonas reinhardtii*.[10,25,40] Our results support the conclusion that *Cf* sGC1 is activated in a similar manner to α/β-type sGCs. sGCs are obligate dimeric proteins. In the heterodimeric CAT domain, neither the α nor the β subunit CAT domains contain the full set of residues necessary for catalysis.[20,41,42] Therefore, α and β subunits must complement each other to form a complete catalytic active site. In *Cf* sGC1, each subunit contains all necessary residues for catalysis.[37] Based on this, *Cf* sGC1 was predicted to be active as a homodimer with two catalytic active sites arranged symmetrically in the catalytic domain. As predicted, *Cf* sGC1 is active as a homodimer, as we previous showed.[17] Based on available transcriptomic data, *Cf* sGC1 is the only predicted NO-specific sGC expressed by *C. flexa*.[16] Whether *Cf* sGC1 can form heterodimers with the other sGCs under physiological condition remains to be determined.

α/β-type sGCs are activated by NO binding, and O_2-binding homodimeric sGCs are typically inhibited when a gas ligand binds to the heme cofactor. In α/β-type sGC, the protein is only fully activated when an excess of NO with respect to heme is present. Stone and coworkers showed that CO activates bovine sGC, a prototypical α/β-type sGC, to levels similar to the 1-NO state.[38] This suggests that under 1-NO and CO-bound states, sGC undergoes similar conformational change, despite the different ligation states of the heme cofactor. In *Cf* sGC1, when an excess of CO relative to heme is present, the enzyme reaches an activity level close to the 1-NO state. This is an additional point of similarity between *Cf* sGC1 and a α/β-type sGC, and this distinguishes it from the *Chlamydomonas*-derived sGC Cyg-11 which is more responsive to CO compared to NO.

During the activation of *Cf* sGC1, the protein exhibits distinct kinetics in the three activity states. While the K_M remained similar in the three activity states, the V_{max} was increased approximately 2-fold in each of the steps, indicating that NO is a V-type activator of *Cf* sGC1. This is a distinction from α/β-type sGC, where NO is a mixed-type activator. For rat sGC, the basal to 1-NO state transition was accompanied by a ~ 7-fold reduction in K_M and a 5-fold increase in V_{max}.

The 1-NO to xsNO state transition showed K_M holding constant while the V_{max} increased 60-fold.[43] When the kinetics of the sGC1-CAT truncation were examined, we observed that both the full-length protein under xsNO state and the sGC1-CAT fitted well to the Michaelis-Menten equation. This suggests the two active site pockets of *Cf* sGC1 do not exhibit significant cooperativity in turning over GTP as a substrate. The dissociation constant K_d for sGC1-CAT was in the low micromolar, compared to the estimated nanomolar affinity of full length sGC. This supports a role for the N-terminal domains in maintaining dimerization.

As α/β-type sGCs are obligate heterodimers, they possess a structural asymmetry that is now well appreciated through full-length structural studies that have been reported.[25,40,44] One observation is that sGC heterodimers bind heme only in the β subunit, and the α subunit is unable to bind heme, partly due to the steric hindrance by the N-terminal loop and partly due to the incomplete set of heme ligating residues.[45] Homodimeric sGCs, on the other hand, have two complete heme binding sites and therefore were expected to bind two heme cofactors per homodimer. However, we find that *Cf* sGC1 binds, on average, one heme per homodimer. The observation that a homodimeric sGC binds one heme per dimer has been observed previously, in both the *D. melanogaster* Gyc-88E and the *C. reinhardtii* Cyg-11.[6,10] Heme reconstitution of sGC typically fails to increase heme incorporation at the physiologically relevant binding pocket and most often results in non-specific binding of heme. In this work, non-specific binding of heme was also observed. Stuehr and colleagues proposed that heme insertion into sGC *in vivo* was facilitated by several molecular chaperones, namely Hsp90 and GAPDH, and heme incorporation likely accompanied dimerization as well as large-scale structural rearrangement.[46,47] This requirement is likely relevant even when the protein is expressed in an *E. coli* heterologous host, although the interaction partners in bacteria remain to be determined. *In vitro* reconstitution experiments are unlikely to accurately recapitulate the physiologically relevant heme insertion process.

In α/β-type sGCs, NO and/or small-molecule stimulators induce an extending motion of the PAS and coiled-coil domains of sGC that is then propagated through the coiled-coil into the CAT domain. In the basal activity state, the catalytic active site is sterically hindered from binding substrate, but upon activation, the site becomes more accessible. Structural asymmetry is also notable during the hinge-like conformational transition between the bent, basal activity state and the more open, fully active state. In the basal activity state, the contracted conformation of the sGC has the α H-NOX domain further distal to the C-terminus than the β H-NOX domain. As noted above, *Cf* sGC1 is a homodimer, and thus does not have interactions between the CAT domain and the H-NOX domains of either subunit that can be differentiated based on sequence. Unsurprisingly, AlphaFold 3 predicts the sGC1 dimer to form a symmetric conformation with relative positions of the domains analogous to the extended state of α/β-type sGC. The AlphaFold prediction does not provide any information on potential alternative conformations, as all 5 ranked predictions converged on the same architecture.

The SAXS studies reported in this work suggest that in solution, the conformation of *Cf* sGC1 is more akin to the compact, asymmetric α/β-type sGC rather than the extended, symmetric AlphaFold prediction. For both NO-present *Cf* sGC1 and NO-present *Ms* sGC, while excess NO was included in the SAXS conditions, the experimental SAXS profile resembles, but did not fully match the expected SAXS profile of a fully active sGC state.[25] In NO-free conditions, both proteins adopt a relatively compact structure. Upon addition of excess NO, the SAXS profiles indicate a

subtle opening, but not to a fully extended state. In the case of *Ms* sGC, Horst et al. interpreted the SAXS profile to arise from a mixture of protein in basal and fully active conformations.[25] The inline UV-vis setup used in our experiment allowed for single UV-vis spectrum to be collected. Based on the UV-vis spectra collected after the elution, we concluded that *Cf* sGC1 remained in the fully NO-bound state as it passed through the X-ray cell. Given the concentration of NO added, it is likely that *Cf* sGC1 remained in the presence of excess NO throughout the experiment. Given this, it is likely that additional factors, such as presence of substrate may be required to shift the conformational equilibrium towards the fully active conformation. Moreover, both DENSS analysis and the comparison to *Ms* sGC SAXS profiles would suggest that *Cf* sGC1 can adopt separate asymmetric conformations under varying NO conditions. The cause and nature of *Cf* sGC1 structural asymmetry remains to be explored at higher resolution. However, we predict that structural asymmetry is closely tied to the asymmetry in heme incorporation. Heme binding to one of the H-NOX domains of *Cf* sGC1 may trigger local shifts in protein folding that create favorable interaction sites with the CC domain and introduce structural asymmetry.

While *Cf* sGC1 is the first of its kind to be characterized, it is likely not unique. NO production have been discovered in sponge species such as *Spongilla lacustris* and placozoan species *Trichoplax adhaerens*, both belonging to early diverging animal lineages, and for each species NOS and sGC have been identified in their genome.[48,49] Sequence alignment suggests that at least one sGC identified in either *S. lacustris* or *T. adhaerens* lack the distal pocket tyrosine associated with regulation by O_2 binding, and two such sGC in *T. adhaerens* are predicted to be active as homodimers based on their sequence. Our work probing the activity and regulation of *Cf* sGC1 will help expand the knowledge of the biochemical basis of NO signaling to diverse groups of animals.

In conclusion, our study shows that *Cf* sGC1 is a homodimeric sGC with similar biochemical characteristics to a α/β-type sGC. While the mechanism of activation of α/β-type sGC has been better understood owing to the newly available structures in the past five years, some molecular details of this process, especially concerning the role of additional equivalents of NO, remain elusive. Our observations risen from studying *Cf* sGC1 are suggestive of a general scheme of allosteric regulation of activity in both homodimeric and heterodimeric sGCs. Further, since *Cf* sGC1 can be expressed in *E. coli,* it could be a powerful model system for studying the process of sGC activation.

References

(1) Aso, Y.; Ray, R. P.; Long, X.; Bushey, D.; Cichewicz, K.; Ngo, T. T.; Sharp, B.; Christoforou, C.; Hu, A.; Lemire, A.; Tillberg, P.; Hirsh, J.; Litwin-Kumar, A.; Rubin, G. M. Nitric Oxide Acts as a Cotransmitter in a Subset of Dopaminergic Neurons to Diversify Memory Dynamics. *eLife* **2019**, 8:e49257

(2) Denninger, J. W.; Marletta, M. A. Guanylate Cyclase and the ·NO/cGMP Signaling Pathway. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1999**, *1411* (2–3), 334–350

(3) Horst, B. G.; Marletta, M. A. Physiological Activation and Deactivation of Soluble Guanylate Cyclase. *Nitric Oxide* **2018**, *77*, 65–74

(4) Koglin, M.; Behrends, S. Native Human Nitric Oxide Sensitive Guanylyl Cyclase: Purification and Characterization. *Biochem. Pharmacol.* **2004**, *67* (8), 1579–1585

(5) Stone, J. R.; Marletta, M. A. Spectral and Kinetic Studies on the Activation of Soluble Guanylate Cyclase by Nitric Oxide. *Biochemistry* **1996**, *35* (4), 1093–1099

(6) Huang, S. H.; Rio, D. C.; Marletta, M. A. Ligand Binding and Inhibition of an Oxygen-Sensitive Soluble Guanylate Cyclase, Gyc-88E, from *Drosophila*. *Biochemistry* **2007**, *46* (51), 15115–15122

(7) Gray, J. M.; Karow, D. S.; Lu, H.; Chang, A. J.; Chang, J. S.; Ellis, R. E.; Marletta, M. A.; Bargmann, C. I. Oxygen Sensation and Social Feeding Mediated by a *C. elegans* Guanylate Cyclase Homologue. *Nature* **2004**, *430* (6997), 317–322

(8) Nighorn, A.; Byrnes, K. A.; Morton, D. B. Identification and Characterization of a Novel β Subunit of Soluble Guanylyl Cyclase That Is Active in the Absence of a Second Subunit and Is Relatively Insensitive to Nitric Oxide. *J. Biol. Chem.* **1999**, *274* (4), 2525–2531

(9) Morton, D. B. Atypical Soluble Guanylyl Cyclases in Drosophila Can Function as Molecular Oxygen Sensors. *J. Biol. Chem.* **2004**, *279* (49), 50651–50653

(10) Horst, B. G.; Stewart, E. M.; Nazarian, A. A.; Marletta, M. A. Characterization of a Carbon Monoxide-Activated Soluble Guanylate Cyclase from *Chlamydomonas reinhardtii*. *Biochemistry* **2019**, *58* (17), 2250–2259

(11) Zimmer, M.; Gray, J. M.; Pokala, N.; Chang, A. J.; Karow, D. S.; Marletta, Michael. A.; Hudson, M. L.; Morton, D. B.; Chronis, N.; Bargmann, C. I. Neurons Detect Increases and Decreases in Oxygen Levels Using Distinct Guanylate Cyclases. *Neuron* **2009**, *61* (6), 865–879

(12) López-Escardó, D.; Grau-Bové, X.; Guillaumet-Adkins, A.; Gut, M.; Sieracki, M. E.; Ruiz-Trillo, I. Reconstruction of Protein Domain Evolution Using Single-Cell Amplified Genomes of Uncultured Choanoflagellates Sheds Light on the Origin of Animals. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2019**, *374* (1786), 20190088

(13) Richter, D. J.; Fozouni, P.; Eisen, M. B.; King, N. Gene Family Innovation, Conservation and Loss on the Animal Stem Lineage. *eLife* **2018**, *7*

(14) Hoffmeyer, T. T.; Burkhardt, P. Choanoflagellate Models — *Monosiga brevicollis* and *Salpingoeca rosetta*. *Curr. Opin. Genet. Dev.* **2016**, *39*, 42–47

(15) Goldstein, B.; King, N. The Future of Cell Biology: Emerging Model Organisms. *Trends. Cell. Biol.* **2016**, *26* (11), 818–824

(16) Brunet, T.; Larson, B. T.; Linden, T. A.; Vermeij, M. J. A.; McDonald, K.; King, N. Light-Regulated Collective Contractility in a Multicellular Choanoflagellate. *Science* **2019**, *366* (6463), 326–334

(17) Reyes-Rivera, J.; Wu, Y.; Guthrie, B. G. H.; Marletta, M. A.; King, N.; Brunet, T. Nitric Oxide Signaling Controls Collective Contractions in a Colonial Choanoflagellate. *Curr. Biol.* **2022**, *32* (11), 2539-2547.e5

(18) Zhao, Y.; Brandish, P. E.; DiValentin, M.; Schelvis, J. P. M.; Babcock, G. T.; Marletta, M. A. Inhibition of Soluble Guanylate Cyclase by ODQ. *Biochemistry* **2000**, *39* (35), 10848–10854

(19) Boon, E. M.; Huang, S. H.; Marletta, M. A. A Molecular Basis for NO Selectivity in Soluble Guanylate Cyclase. *Nat. Chem. Biol.* **2005**, *1* (1), 53–59

(20) Winger, J. A.; Derbyshire, E. R.; Lamers, M. H.; Marletta, M. A.; Kuriyan, J. The Crystal Structure of the Catalytic Domain of a Eukaryotic Guanylate Cyclase. *BMC Struct. Biol.* **2008**, *8* (1), 42

(21) Acerbo, A. S.; Cook, M. J.; Gillilan, R. E. Upgrade of MacCHESS Facility for X-Ray Scattering of Biological Macromolecules in Solution. *J. Synchrotron Radiat.* **2015**, *22* (1), 180–186

(22) Skou, S.; Gillilan, R. E.; Ando, N. Synchrotron-Based Small-Angle X-Ray Scattering of Proteins in Solution. *Nat. Protoc.* **2014**, *9* (7), 1727–1739

(23) Illava, G.; Gillilan, R.; Ando, N. Development of In-Line Anoxic Small-Angle X-Ray Scattering and Structural Characterization of an Oxygen-Sensing Transcriptional Regulator. *J. Biol. Chem.* **2023**, *299* (8), 105039

(24) Maragos, C. M.; Morley, D.; Wink, D. A.; Dunams, T. M.; Saavedra, J. E.; Hoffman, A.; Bove, A. A.; Isaac, L.; Hrabie, J. A.; Keefer, L. K. Complexes of · NO with Nucleophiles as Agents for the Controlled Biological Release of Nitric Oxide. Vasorelaxant Effects. *J. Med. Chem.* **1991**, *34* (11), 3242–3247

(25) Horst, B. G.; Yokom, A. L.; Rosenberg, D. J.; Morris, K. L.; Hammel, M.; Hurley, J. H.; Marletta, M. A. Allosteric Activation of the Nitric Oxide Receptor Soluble Guanylate Cyclase Mapped by Cryo-Electron Microscopy. *eLife* **2019**, *8*

(26) Hopkins, J. B. *BioXTAS RAW 2* : New Developments for a Free Open-Source Program for Small-Angle Scattering Data Reduction and Analysis. *J. Appl. Crystallogr.* **2024**, *57* (1), 194–208

(27) Hopkins, J. B.; Gillilan, R. E.; Skou, S. *BioXTAS RAW* : Improvements to a Free Open-Source Program for Small-Angle X-Ray Scattering Data Reduction and Analysis. *J. Appl. Crystallogr.* **2017**, *50* (5), 1545–1553

(28) Meisburger, S. P.; Taylor, A. B.; Khan, C. A.; Zhang, S.; Fitzpatrick, P. F.; Ando, N. Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-Ray Scattering. *J. Am. Chem. Soc.* **2016**, *138* (20), 6506–6516

(29) Hansen, S. Bayesian Estimation of Hyperparameters for Indirect Fourier Transformation in Small-Angle Scattering. *J. Appl. Crystallogr.* **2000**, *33* (6), 1415–1421

(30) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. FoXS, FoXSDock and MultiFoXS: Single-State and Multi-State Structural Modeling of Proteins and Their Complexes Based on SAXS Profiles. *Nucleic Acids Res.* **2016**, *44* (W1), W424–W429

(31) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. Accurate SAXS Profile Computation and Its Assessment by Contrast Variation Experiments. *Biophys. J.* **2013**, *105* (4), 962–974

(32) Manalastas-Cantos, K.; Konarev, P. V.; Hajizadeh, N. R.; Kikhney, A. G.; Petoukhov, M. V.; Molodenskiy, D. S.; Panjkovich, A.; Mertens, H. D. T.; Gruzinov, A.; Borges, C.; Jeffries, C. M.; Svergun, D. I.; Franke, D. *ATSAS 3.0* : Expanded Functionality and New Tools for Small-Angle Scattering Data Analysis. *J. Appl. Crystallogr.* **2021**, *54* (1), 343–355

(33) Grant, T. D. Ab Initio Electron Density Determination Directly from Solution Scattering Data. *Nat. Methods* **2018**, *15* (3), 191–193

(34) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Žemgulytė, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M. R.; Perlin, K.; Potapenko, A.; Savy, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Žídek, A.; Bapst, V.; Kohli, P.; Jaderberg, M.; Hassabis, D.; Jumper, J. M. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, *630* (8016), 493–500

(35) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46* (W1), W296–W303

(36) Liu, R.; Kang, Y.; Chen, L. NO Binds to the Distal Site of Haem in the Fully Activated Soluble Guanylate Cyclase. *Nitric Oxide* **2023**, *134–135*, 17–22

(37) Barr, I.; Guo, F. Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *Bio. Protoc.* **2015**, *5* (18)

(38) Stone, J. R.; Marletta, M. A. Soluble Guanylate Cyclase from Bovine Lung: Activation with Nitric Oxide and Carbon Monoxide and Spectral Characterization of the Ferrous and Ferric States. *Biochemistry* **1994**, *33* (18), 5636–5640

(39) Winger, J. A.; Marletta, M. A. Expression and Characterization of the Catalytic Domains of Soluble Guanylate Cyclase: Interaction with the Heme Domain. *Biochemistry* **2005**, *44* (10), 4083–4090

(40) Kang, Y.; Liu, R.; Wu, J.-X.; Chen, L. Structural Insights into the Mechanism of Human Soluble Guanylate Cyclase. *Nature* **2019**, *574* (7777), 206–210

(41) Allerston, C. K.; von Delft, F.; Gileadi, O. Crystal Structures of the Catalytic Domain of Human Soluble Guanylate Cyclase. *PLoS One* **2013**, *8* (3), e57644

(42) Seeger, F.; Quintyn, R.; Tanimoto, A.; Williams, G. J.; Tainer, J. A.; Wysocki, V. H.; Garcin, E. D. Interfacial Residues Promote an Optimal Alignment of the Catalytic Center in Human

Soluble Guanylate Cyclase: Heterodimerization Is Required but Not Sufficient for Activity. *Biochemistry* **2014**, *53* (13), 2153–2165

(43) Fernhoff, N. B.; Derbyshire, E. R.; Marletta, M. A. A Nitric Oxide/Cysteine Interaction Mediates the Activation of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U S A* **2009**, *106* (51), 21602–21607

(44) Liu, R.; Kang, Y.; Chen, L. Activation Mechanism of Human Soluble Guanylate Cyclase by Stimulators and Activators. *Nat. Commun.* **2021**, *12* (1), 5492

(45) Wittenborn, E. C.; Marletta, M. A. Structural Perspectives on the Mechanism of Soluble Guanylate Cyclase Activation. *Int. J. Mol. Sci.* **2021**, *22* (11), 5439

(46) Ghosh, A.; Stuehr, D. J. Soluble Guanylyl Cyclase Requires Heat Shock Protein 90 for Heme Insertion during Maturation of the NO-Active Enzyme. *Proc. Natl. Acad. Sci. U S A* **2012**, *109* (32), 12998–13003

(47) Ghosh, A.; Stasch, J. P.; Papapetropoulos, A.; Stuehr, D. J. Nitric Oxide and Heat Shock Protein 90 Activate Soluble Guanylate Cyclase by Driving Rapid Change in Its Subunit Interactions and Heme Content. *J. Biol. Chem.* **2014**, *289* (22), 15259–15271

(48) Moroz, L. L.; Romanova, D. Y.; Nikitin, M. A.; Sohn, D.; Kohn, A. B.; Neveu, E.; Varoqueaux, F.; Fasshauer, D. The Diversification and Lineage-Specific Expansion of Nitric Oxide Signaling in Placozoa: Insights in the Evolution of Gaseous Transmission. *Sci. Rep.* **2020**, *10* (1), 13020

(49) Ruperti, F.; Becher, I.; Stokkermans, A.; Wang, L.; Marschlich, N.; Potel, C.; Maus, E.; Stein, F.; Drotleff, B.; Schippers, K. J.; Nickel, M.; Prevedel, R.; Musser, J. M.; Savitski, M. M.; Arendt, D. Molecular Profiling of Sponge Deflation Reveals an Ancient Relaxant-Inflammatory Response. *Curr. Biol.* **2024**, *34* (2), 361-375.e9

Chapter 4
Probing non-heme NO interaction of *Cf* sGC1

Summary

The soluble guanylate cyclase *Cf* sGC1 from the choanoflagellate species *Choanoeca flexa* is a highly specific sensor of NO. *Cf* sGC1 exhibits a three-state activation profile *in vitro* reminiscent of α/β-type sGC found in mammals and insects, where NO binding to the heme is not sufficient to fully activate *Cf* sGC1. The full activation of α/β-type sGC has been proposed to be regulated through non-heme NO interactions, and cysteine residues may be involved. To probe the mechanism of non-heme NO interaction in regulating *Cf* sGC1, cysteine variants at conserved sites (C123, C355 and C489) were prepared and characterized. Cysteine variants of *Cf* sGC1 displayed similar gas ligand binding properties to the wild type (WT), showing a 5-coordinate unliganded Fe(II) heme with an absorbance maximum at 428 nm, and 5-coordinate Fe(II)-NO complex with an absorbance maximum at 399 nm. The cysteine variants did not perturb the NO-mediated activation of *Cf* sGC1. Furthermore, methyl methanethiosulfonate (MMTS), a cysteine-specific labeling reagent that inhibits sGC activity was tested against *Cf* sGC1. The effects of MMTS were hypothesized to be disruption of cysteine-NO interactions, or allosteric blocking the conformational change that takes place in sGC activation. The cysteine variants did not display a significantly different response to MMTS labeling compared to WT. Taken together, these data suggest that C123, C355 and C489 are likely not independently responsible for regulating non-heme NO interaction in *Cf* sGC1, nor are they involved in MMTS-mediated inhibition of *Cf* sGC1.

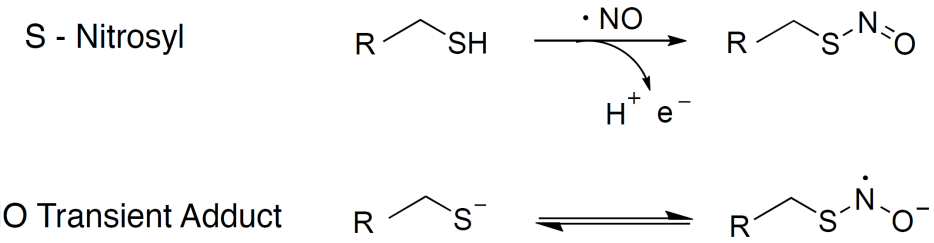
Introduction

In mammals, soluble guanylate cyclases (sGC) are heme-containing nitric oxide sensing enzymes that play important roles in both the vasculature and the neuronal systems.[1–3] The activation of sGC requires NO binding to the heme in the N-terminal heme-nitric oxide oxygen-binding domain (H-NOX), which triggers a large-scale conformational shift that leads to increased catalytic efficiency of the catalytic (CAT) domain.[4,5] The high-resolution structures that recently became available have provided many residue-level interactions that contribute to the regulation of sGC activity can be studied in detail.[4–7] However, the high-resolution structures fell short of explaining some crucial biochemical details of the activation of sGC. One such detail is the nature of the 1-NO state (NO bound only to the heme) to xsNO state transition during sGC activation.

The activation of heterodimeric α/β-type sGC by NO *in vitro* occurs in several steps. First, NO binds to the heme to form a transient 6-coordinate Fe(II)-NO complex. This complex converts rapidly into a 5-coordinate Fe(II)-NO complex, accompanied by the rupture of the proximal His-Fe bond. In this state (dubbed the 1-NO state), sGC is activated several-fold over basal activity, roughly 15% of full activity.[8,9] sGC in the 5-coordinate Fe(II)-NO bound state is sensitive to additional equivalents of NO. When NO is supplied in excess of that bound to the heme, sGC becomes fully active, a state that is referred to as the xsNO state. Removal of excess NO through gel filtration or by adding a NO-specific trap returns the enzyme to the 15% activity state without loss of the heme-bound NO.[8] The physiological importance of the interaction with excess NO in sGC can be further illustrated in terms of sGC inactivation. NO is essentially irreversibly bound

to the heme cofactor of sGC with a $K_d < 1.2 \times 10^{-12}$ M.[10] Dissociation of NO from the 5-coordinate Fe(II) heme-NO complex requires rebinding of the axial histidine and has a rate constant on the order of 1×10^{-4} s^{-1}.[11] This rate is much slower than the physiological timescale of sGC inactivation, which is in the seconds range.[8,12] Therefore, the NO-heme ligated state of sGC is likely the relevant basal state *in vivo.*

Given that the α/β-type sGC requires excess equivalents of NO to reach full activity, the mechanism of excess NO action was called into question. The UV-vis absorption spectrum of α/β-type sGC in the 1-NO state is indistinguishable from the xsNO state, suggesting that the heme remained in the 5-coordinate state during the activation of sGC and one implication is that there is a second, non heme binding site for NO.[8] The hypothesis that sGC interacts with NO at a non-heme site was further supported by the work of Derbyshire and colleagues. When a non-physiological ligand (butyl isocyanide, BIC) was used to occupy the heme site, supplementation of NO resulted in an 18-fold sGC activation compared to BIC only. UV-vis spectroscopy indicated that BIC remained bound in a 6-coordinate state throughout the duration of the experiment, suggesting that the effects of NO supplementation did not occur at the heme site..[13]



Scheme 3.1. Comparison between S-nitrosation and the proposed radical anion S-NO transient adduct. Formation of S-nitrosyl requires a 1-electron reduction, while the radical anion does not require redox chemistry and is freely reversible.

A hypothesis regarding the chemical property of the non-heme NO binding site was presented by further experiments carried out by Fernhoff and colleagues. *Rattus norvegicus* sGC was treated with the cysteine-active labeling reagent methyl methanethiosulfonate (MMTS). Post treatment, sGC exhibits 1-NO like activity despite the presence of excess NO.[14] The MMTS-treated sGC showing 1-NO activity level can be further activated using small molecule stimulator drugs such as YC-1. MMTS-treated sGC can be rescued to show near full activation in the presence of excess NO by DTT treatment to remove S-methyl adduct through reduction. Fernhoff and coworkers concluded that cysteines are responsible for non-heme NO interaction in sGC, and MMTS blocks this interaction from taking place. Given that experiments above are carried out under oxygen-free environments and using protein with a reduced heme cofactor, a direct interaction between the cysteine thiolate and the NO to form a transient radical anion species is more likely than S-nitrosation, as it does not require redox chemistry to take place and furthermore, is freely reversible whereas the S-nitrosated sGC would not be readily reversible (Scheme 3.1). However, the exact cysteine residues that are involved in this proposed pathway have not been identified. The large conformational change that sGC undergoes during activation suggests a second hypothesis explaining sGC inhibition by MMTS. The S-methyl group, though small and hydrophobic, may still block the conformational change of a sGC during activation. The two

pathways described above may both be at play for preventing sGC activation and can be tested independently through cysteine mutagenesis.

In Chapter 2 and 3, a homodimeric sGC, *Cf* sGC1 that exhibits a 3-stage activation profile in response to NO binding was characterized. We hypothesized that full activation of *Cf* sGC1 is driven by a similar mechanism of non-heme NO interaction. Sequence alignment of *Cf* sGC1 against prototypical α/β-type sGC revealed only three conserved cysteine residues, and variants of *Cf* sGC1 were prepared to interrogate the role of conserved cysteines in regulating the activation of *Cf* sGC1. Activity assays of *Cf* sGC1 variants indicate that the conserved cysteine residues were not responsible for mediating non-heme NO interaction, and MMTS labeling studies suggest that the conserved cysteines were not responsible for MMTS-mediated inhibition of *Cf* sGC1 activity.

Methods

Materials

Primers were purchased from Integrated DNA Technologies. Gibson Master Mix and *E. coli* competent cell stocks were purchased from the UC Berkeley QB3 MacroLab. Phusion Polymerase and KLD enzyme kit were purchased from New England Biolabs. Plasmid extraction and DNA purification kits were purchased from Qiagen. Terrific Broth media powder, kanamycin, benzamidine and 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) were purchased from Research Products International. Sodium dithionite ($Na_2S_2O_4$), DNAse I from bovine pancreas, sodium phosphate, sodium acetate, magnesium acetate, methyl methanethiosulfonate (MMTS) and guanosine 5'-triphosphate sodium salt were purchased from Sigma Aldrich. Isopropyl β-D-1-thiogalactopyranoside (IPTG), 4-(2-hydroxyethyl)piperazine ethanesulfonic acid (HEPES), Pierce BCA Protein Assay Kit, ammonium acetate and acetonitrile (HPLC grade) were purchased from Thermo Fisher Scientific. Lysozyme from chicken egg white and sodium chloride were purchased from VWR Scientific. Dithiothreitol (DTT) was purchased from BACHEM. Imidazole was purchased from Oakwood Chemical. Vivaspin spin concentrators were purchased from Sartorius. D-Aminolevulinic acid (d-ALA) and hemin chloride were purchased from Spectrum Chemical Company. proliNONOate was purchased from Cayman Chemical Company.

Construction of variants

Variants were constructed based on the pET_*Cf* sGC1 construct described in chapter 3. Site directed mutagenesis was carried out by first amplifying the DNA fragment using mutagenic primers, and then ligated using the KLD reaction mix following manufacturer's instructions. The identity of purified plasmid DNA was verified by sequencing at the UC Berkeley Sequencing Facility.

Protein expression and purification

Variant constructs of *Cf* sGC1 were transformed into *E. coli* BL21Star (DE3) competent cells. Overnight culture was prepared by inoculating one colony in 50 mL of Luria Broth media supplemented with 50 µg/mL kanamycin. After overnight shaking at 37 °C, the overnight culture was diluted 1:200 into Terrific Broth (TB) media supplemented with 50 µg/mL kanamycin and 2% glucose. The overnight culture and the large culture were supplemented with 500 µM $FeCl_3$. The large culture in TB media was incubated at 37 °C until cell density reached OD_{600} = 0.6 – 0.8

before protein production was induced by the addition of 500 μM IPTG. d-ALA (1 mM) was supplemented when the cell density of the culture reached OD_{600} = 0.4 – 0.6 and the culture was incubated at 37 °C for an additional 15 minutes before protein production was induced. After induction, the culture was incubated at 18 °C for 16 – 20 hours, and the cells were harvested through centrifugation at 4500 x g, 4 °C for 25 min. The cell paste was collected, flash frozen in liquid nitrogen, and transferred to -80 °C for storage.

Protein purification was carried out at 4 °C unless otherwise specified. Cell pellets were thawed in an ice water bath and resuspended in ice cold Buffer A (50 mM sodium phosphate, 150 mM NaCl, 5 mM imidazole, 5 % glycerol, pH 8.0) supplemented with 50 mM benzamidine, 0.2 mM AEBSF, 0.25 mg/mL DNAse I, and 0.25 mg/mL lysozyme. The cell resuspension was lysed using a high pressure homogenizer (Avestin Emulsiflex C5). The resulting cell lysate was clarified by centrifugation (36,000 x g, 55 min) and passed through a gravity column loaded with His60 Ni Superflow Resin equilibrated with Buffer A. The column was subsequently washed with 10 CV Buffer A, 10 CV Buffer B (50 mM sodium phosphate, 200 mM NaCl, 40 mM imidazole, 5 % glycerol, pH 8.0) and His-tagged protein was eluted with 5 CV Buffer C (50 mM sodium phosphate, 200 mM NaCl, 400 mM imidazole, 5 % glycerol, pH 8.0). Protein-containing fractions determined by SDS-PAGE electrophoresis were combined and concentrated using a 50 kDa molecular weight cut-off spin concentrator and diluted 10x in Buffer D (25 mM triethanolamine, 25 mM NaCl, 5 mM DTT, pH 7.4). The diluted protein was loaded on to a POROS HQ20 anion exchange column (Thermo Scientific) equilibrated with Buffer D and the protein was eluted over a gradient of 5% to 50% Buffer E (25 mM triethanolamine, 750 mM NaCl, 5 mM DTT, pH 7.4) over 17 CV at 5 mL/min. Fractions with a ratio of A_{428}:A_{280} > 1 were combined, concentrated to 25 – 50 μM and flash frozen in liquid nitrogen. Protein aliquots were stored at -80 °C for future use.

Activity assay

Initial rates of *Cf* sGC1 variants were measured by quantifying the production of cGMP in endpoint assays performed at 25 °C under anoxic conditions. Prior to the assay, *Cf* sGC1 variants were reduced using 5 mM $Na_2S_2O_4$ and desalted into Ar-sparged Buffer F (50 mM HEPES, 150 mM NaCl, 5 % glycerol, pH 7.5, 0.22 μm filtered) using a pre-equilibrated 7.5 kDa MWCO Zeba Spin Desalting column (Thermo Scientific). Concentration of the protein was determined by the absorbance at 428 nm (ε = 101,000 M^{-1} cm^{-1}).[15] xsNO protein sample was prepared by the addition of 250 μM proliNONOate. The sample was incubated for 5 minutes at 4 °C, then desalted into Buffer F using a preequilibrated spin desalting column to yield the 1-NO protein sample. For MMTS-labeled protein samples, a 50 mM stock solution of MMTS was diluted to specific concentrations ranging from 1 to 10 μM as needed to label 1 μM *Cf* sGC1 WT and variant protein. Labeling reaction mixtures were incubated at 4 °C for 30 minutes. Activity assay mixtures were prepared in Ar-sparged Buffer F supplemented with 5 mM DTT and $MgCl_2$ greater or equal to two-fold the concentration of GTP and included 50 nM wild type and variant *Cf* sGC1 protein. For the xsNO sample, NO was supplemented by the addition of 50 μM proliNONOate. Reaction was initiated by the addition of GTP and aliquots of the reaction mixture were removed and quenched using 1 % formic acid. The results of the enzymatic reaction were analyzed using reverse phase HPLC as previously reported.[16] Briefly, the reaction timepoint samples were injected onto an Eclipse Plus C18 column (4.6 x 100 mm, 3.5 μm particle size, Agilent). A gradient composed of 20 mM ammonium acetate, 0.1 % formic acid (Buffer G) and 99.9 % acetonitrile, 0.1 % formic acid (Buffer H) was used to separate GTP and cGMP. The gradient was as follows: 0 – 6 min, 2%

Buffer H; 6 – 7.5 min, 25 – 100% Buffer H; 7.5 – 8.5 min, 100 % Buffer H. Concentrations of cGMP were determined from the peak area eluting at 4.2 minutes using a standard curve constructed from 1.25 – 20 μM cGMP standards.

Mass spectrometry

MMTS-labeled protein was characterized using mass spectrometry. Prior to injection, samples were centrifuged at 4 °C, 21,130 xg for 10 minutes, and the supernatant was collected for analysis. Liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) was carried out on an Agilent 1200 Series LC system coupled to an Agilent 6224 time-of-flight mass spectrometer in positive ion mode. Chromatographic separation of the protein sample was achieved over a Proswift RP-4H (1.0 mm × 50 mm, monolithic phenyl, Dionex) column using a mobile phase composed of 99.9% water, 0.1% formic acid v/v (MS solvent A) and 99.9% acetonitrile, 0.1% formic acid v/v (MS solvent B). The elution gradient was developed over 8 minutes, 5% to 100% B at 0.3 mL/min. Data were collected and deconvoluted using Agilent MassHunter software. Mass spectra were visualized using the open-source online tool Chartograph (chartograph.com).

Results and Discussions

Given the importance of non-heme NO interaction with sGC, and the hypothesis that cysteines play a crucial role in regulating sGC activity, we decided to further probe the mechanism of *Cf* sGC1 activation to understand whether cysteines are involved and explain why *Cf* sGC1 shows a similar activation profile to α/β-type sGC. If cysteine thiolate interacts with NO to form a transient radical anion adduct, this species is expected to either cause a conformational change, or to change local amino acid contacts for sGC to better access the fully active conformation. If this hypothesis was true, then such cysteines are likely highly conserved on sGCs that show the three-stage activation profile. As such, we decided to create cysteine variants of *Cf* sGC1 for biochemical characterization. Given that *Cf* sGC1 can be overexpressed in *E. coli*, biochemical characterization of purified variant proteins is made possible by a comparatively simple purification process and much improved yield compared to insect cell expression and purification.

Conservation of cysteines across α/β-type sGCs and Cf sGC1

Prior to the construction of the variants, sequence alignment and sequence similarity network tools were employed for validating the conservation of cysteines at positions of interest. Prior to the discovery of *Cf* sGC1, sequence alignment had been used to narrow down the list of conserved cysteines within α/β-type sGCs.[17] When *Hs* sGC was aligned against *Ms* sGC, only 16 total cysteines were conserved. However, when *Cf* sGC1 was aligned to *Rn* sGC and *Ms* sGC, the list of conserved cysteines was narrowed down to only three. Based on *Cf* sGC1 numbering, the three cysteines were at positions 123, 355 and 489. The three cysteines are situated in the H-NOX domain, PAS domain and N-terminus of the CAT domain, respectively. A fourth cysteine, C177 (*Cf* numbering), is located on the same loop connecting β strands 3 and 4 of the H-NOX compared to β subunit C173 (*Rn* numbering). However, the sequence of *Cf* sGC1 in this region did not match the consensus sequence of sGC β subunits, so C177 was not considered further.

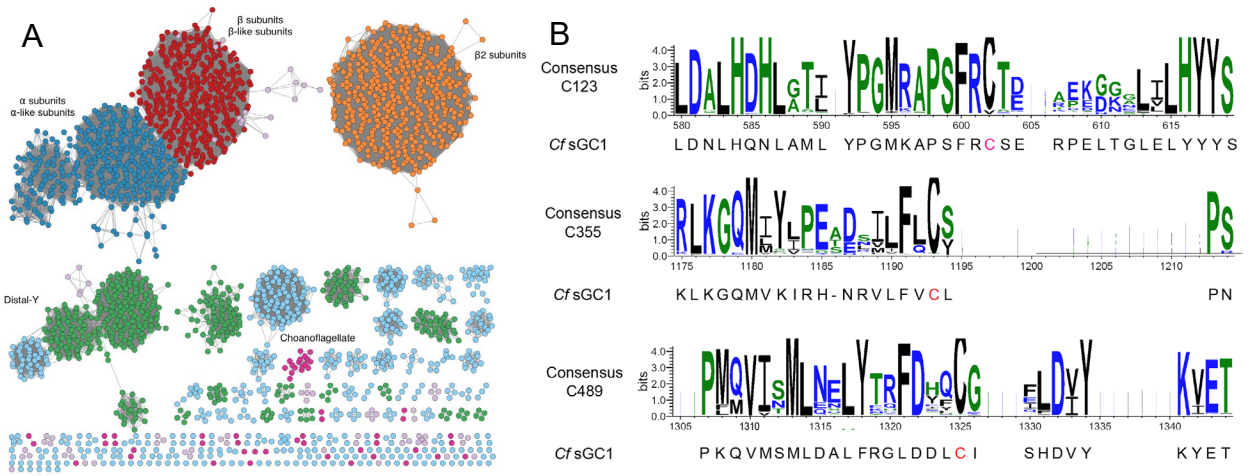


Figure 4.1. Bioinformatics analysis of conserved cysteine positions of *Cf* sGC1. (A) Sequence similarity network of sGC proteins, generated using protein family PF07700 (H-NOX domain-containing proteins) and choanoflagellate sGC sequences. Edges correspond to sequence identity >57 %. α and β subunits were separated into clusters under these parameters. Green-colored clusters contain sGC sequences that contain a conserved distal pocket Tyr residue and are predicted to bind O_2. Choanoflagellate-derived sequences are colored violet, and sequences derived from early diverging animals (placozoa, porifera, cnidaria) are colored light purple. Sequence similarity network was calculated using the EFI Enzyme Similarity Tool and visualized using Cytoscape 3.0. (B) Consensus sequences calculated based on clustered α and β subunit sequences, compared to the sequence of *Cf* sGC1. *Cf* sGC1 has an overall similar sequence to the consensus sequence at the sites of conserved cysteine residues. Sequence alignment was carried out using ClustalOmega and visualized with WebLogo3.

Following the identification of the three conserved cysteines, their locations were examined to verify whether they are situated in well conserved regions of the α/β-type sGC. To compile a set of α and β subunit sequences, the sequence space of sGC was analyzed using sequence similarity network (SSN).[18] In a SSN, sequences are represented as nodes, and if two sequences have an alignment score above a certain threshold, an edge is drawn to connect them. In this manner, nodes representing similar sequences can be grouped into clusters. Proteins in the same cluster likely exhibit similar biochemical functions. The SSN was constructed using sequences from protein family PF07700 that corresponds to H-NOX domain-containing proteins, and SSN edges as shown correspond sequence identity greater than 57%. Under these conditions, we observed that the α and β subunits each grouped into distinct clusters (Figure 4.1). Analysis of the sequences that grouped into these clusters revealed that they covered most bilaterians clades, including mollusks, annelids, vertebrates and arthropods. The sequences obtained from the α cluster and the β cluster were aligned using ClustalOmega, and the alignment result was visualized using WebLogo3 to highlight sequence conservation. The sequence alignment suggests that the three cysteines are situated at well-conserved regions of α/β-type sGC.

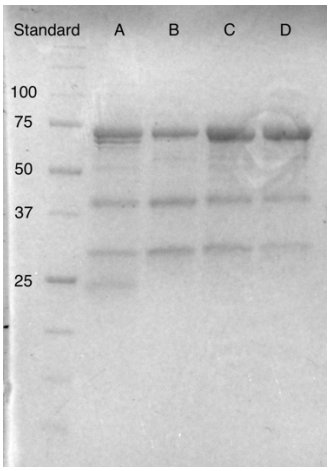


Figure 4.2. Gel electrophoresis of the variants of *Cf* sGC1. Lanes: A, C355A. B, C355G. C, C123S C355G. D, C123S C355G C489A. The variants of *Cf* sGC1 were purified using Ni-affinity chromatography followed by anion exchange chromatography. Minor degradation products were observed in the purified protein and may be a result of reduced stability from mutagenesis. Expected molecular weight for all four variants are around 75 kDa.

Variants of *Cf* sGC1 at positions 123, 355 and 489 were constructed based on the prediction of local solvent accessibility. Conservative single point variants alanine and glycine were constructed at position 355, and a more polar residue serine. A double point variant was constructed for position 123 and 355, with a serine at position 123 and an alanine at position 355. A triple variant was also constructed, with serine at position 123, alanine at 355, and glycine at 489. The double variant and triple variant were both considered conservative mutations so the perturbation to the overall fold of the protein should be minimal. During expression and purification of *Cf* sGC1 variants, the heme cofactor was apparently more readily oxidized, and gel electrophoresis results indicated that the variants were prone to proteolytic degradation (Figure 4.2). This suggests that mutations at the above positions likely caused some decrease in the stability of the protein. After purification, however, the purified protein mostly remained intact and allowed for characterization of catalytic activity.

Activity results of Cf sGC1 variants

Activity assays for variants of *Cf* sGC1 (C355A, C355G. C123S/C355A, C123S/C355A/C489G) were carried out similarly to WT *Cf* sGC1. After the protein was reduced and prior to the assay, gas ligand binding was confirmed by UV-vis spectroscopy. All four variants showed the expected gas ligand properties for a NO-activated sGC, with a 5-coordinate high spin unliganded state and a 5-coordinate high spin NO-bound state (Figure 4.3). This suggests that the point variants did not significantly perturb the heme environment of the protein, despite that C123 is located within the H-NOX domain and close to the heme binding site.

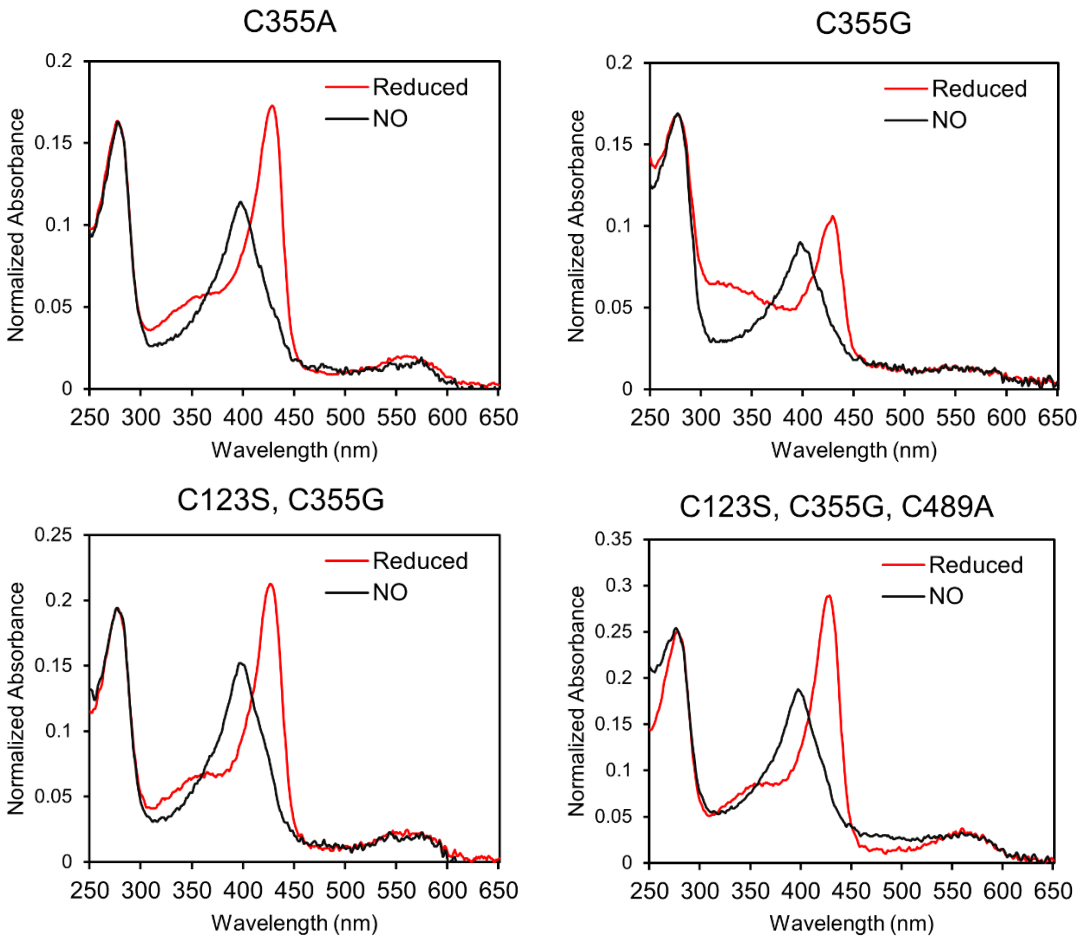


Figure 4.3. UV-vis spectroscopy of the *Cf* sGC1 variants. Under NO-free condition, all four variants show an unliganded heme, with the Soret band absorbing at 428 nm. When NO is added, all four variants form a Fe(II) heme-NO complex, with the Soret band absorbing at 399 nm. No significant difference in ligand binding can be observed for *Cf* sGC1 variants when compared to *Cf* sGC1 WT.

To quickly screen the effects of the mutations on the activity of *Cf* sGC1, activity assays were only done once for each of the variants. While this approach did not allow for establishment of statistical significance of the differences between each ligand bound condition, the same trend can be observed for each variant. Mutations did not affect the three-stage activation profile of *Cf* sGC1; even the triple variant still showed full activation in the presence of NO and a partial activation in the 1-NO state. This data suggests that despite the conservation of the three cysteine residues, they are either not involved in non-heme interaction with NO that triggers the full activation of *Cf* sGC1, or an alternative interaction site can complement the loss of interaction at the site mutated. These results mirror findings previously reported by Horst.[17] Horst cloned cysteine single point mutations of *Manduca sexta* sGC, a prototypical α/β-type sGC, and despite extensive characterization through enriched lysate assays, he was not able to identify a single point mutant that significantly perturbed the full activation of *Ms* sGC by excess NO. This result also helped us rule out one part of the hypothesis, which is that cysteines conserved with α/β-type sGCs are responsible for non-heme NO interaction and for activating the protein to full activity in *Cf* sGC1.

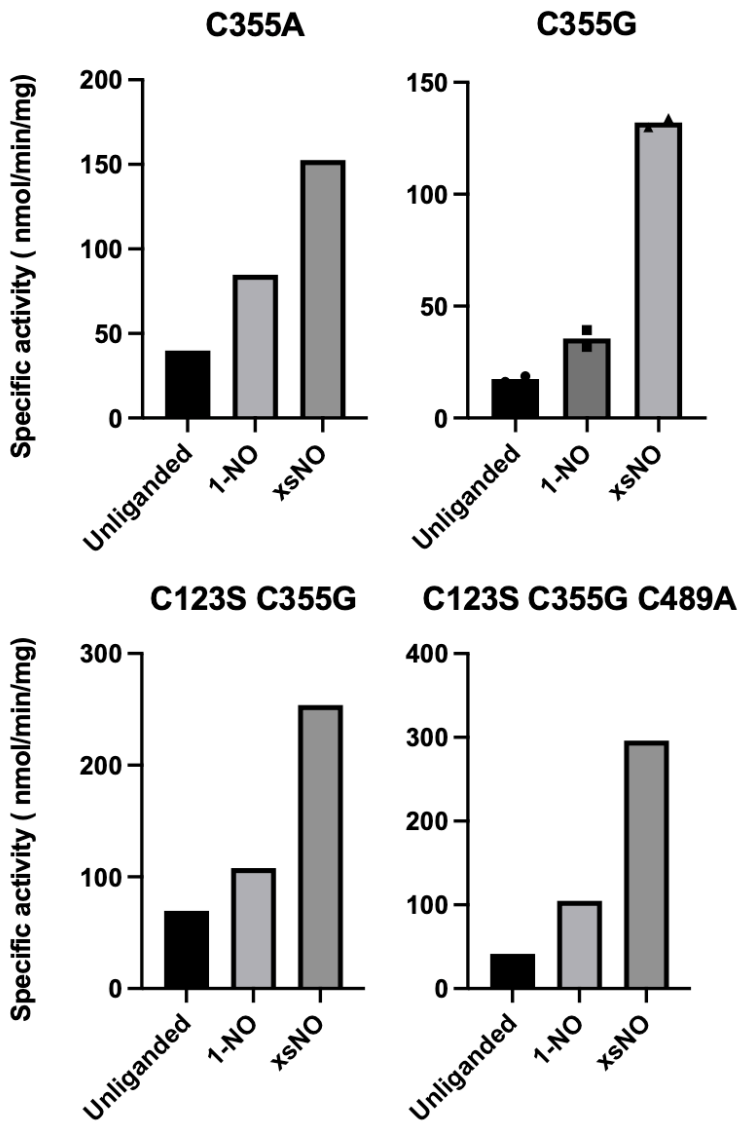


Figure 4.4. Activity of the of *Cf* sGC1 variants. Shown are single activity assay results. Despite minor differences in the absolute value of the activity, the variants did not show in any noticeable change in the respective activity profiles.

MMTS labeling of Cf sGC1 and the variants

Testing the NO-stimulated activity of sGC variants, we ruled out the hypothesis that NO forming transient intermediates at positions 123, 355 or 489 is the trigger for full activation of *Cf* sGC1, and that MMTS works by blocking cysteine-NO interaction at the above positions. Here, the second hypothesis was tested. If the S-methyl adduct from a MMTS labeling experiment could sterically block the protein from accessing the fully active conformation, then by mutating the affected C to a residue that MMTS could not access, we would expect to see the loss of MMTS-mediated inhibition of sGC.

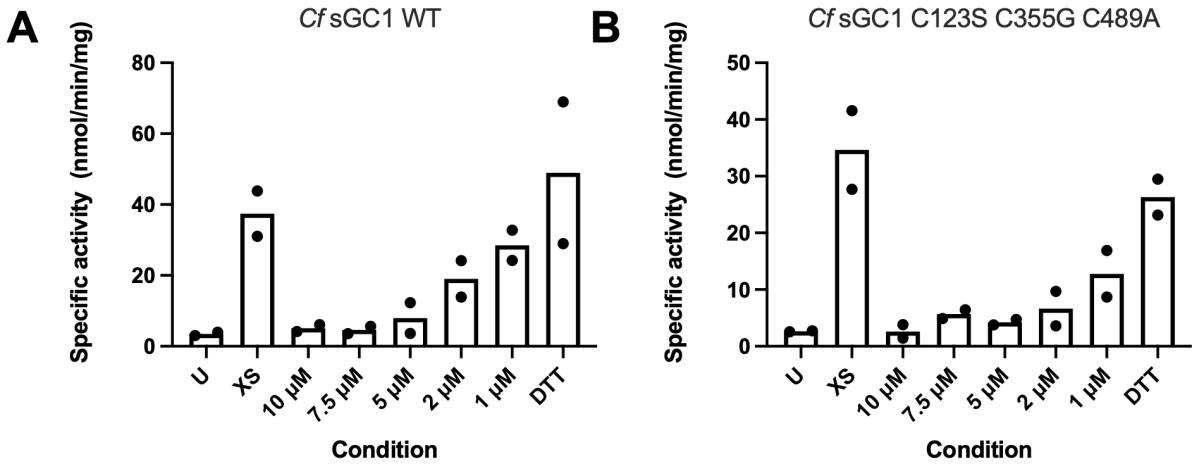


Figure 4.5. Activity of (A) wild type *Cf* sGC1 and (B) triple variant *Cf* sGC1 after MMTS labeling**.** The MMTS titration assay was carried out in duplicate. Comparing the activity of the triple variant *Cf* sGC1 to the activity of *Cf* sGC1 WT after MMTS labeling revealed no significant difference in the response to MMTS inhibition with and without mutation.

MMTS labeling experiments were carried out for *Cf* sGC1 WT protein as well as the triple variant. In the MMTS labeling experiment, 1 µM protein was labeled with 1 – 10 µM MMTS. MMTS labeling can be simply reversed through the addition of DTT, so a control sample with DTT addition was also prepared. For both the WT and the triple variant *Cf* sGC1, inhibition of activity was first observed at ~2 µM, and the protein was inhibited to basal-like activity between 5 and 7.5 µM (Figure 4.5). This result suggests that C123, C355 and C489 were either not accessible to MMTS in the labeling experiment or labeling at these positions did not affect the activity of *Cf* sGC1.

Mass spectrometry of Cf sGC1 labeling products

Given that the MMTS-induced inhibition of sGC occurred in the range of 2 to 5 µM, intact protein mass spectrometry was used to verify that the protein was labeled under the conditions tested, and to quantify the number of labels on the protein after incubation. Intact protein MS results suggest that by using 5 µM MMTS, the majority of *Cf* sGC1 WT and triple variant protein monomer was labeled. The presence of multiple MS peaks corresponding to unmodified protein, +1 S-methyl and +2 S-methyl suggests that the labeling reaction likely did not yield a homogeneous mixture of labeled dimeric *Cf* sGC1. However, our data suggests that most dimeric *Cf* sGC1 after labeling are at least modified at one cysteine position. Compared to *Ms* sGC, a prototypical α/β-type sGC, *Cf* sGC1 is significantly more sensitive to MMTS. The reason for this sensitivity remains to be determined but is likely associated with the location and accessibility of the labeling sites involved.

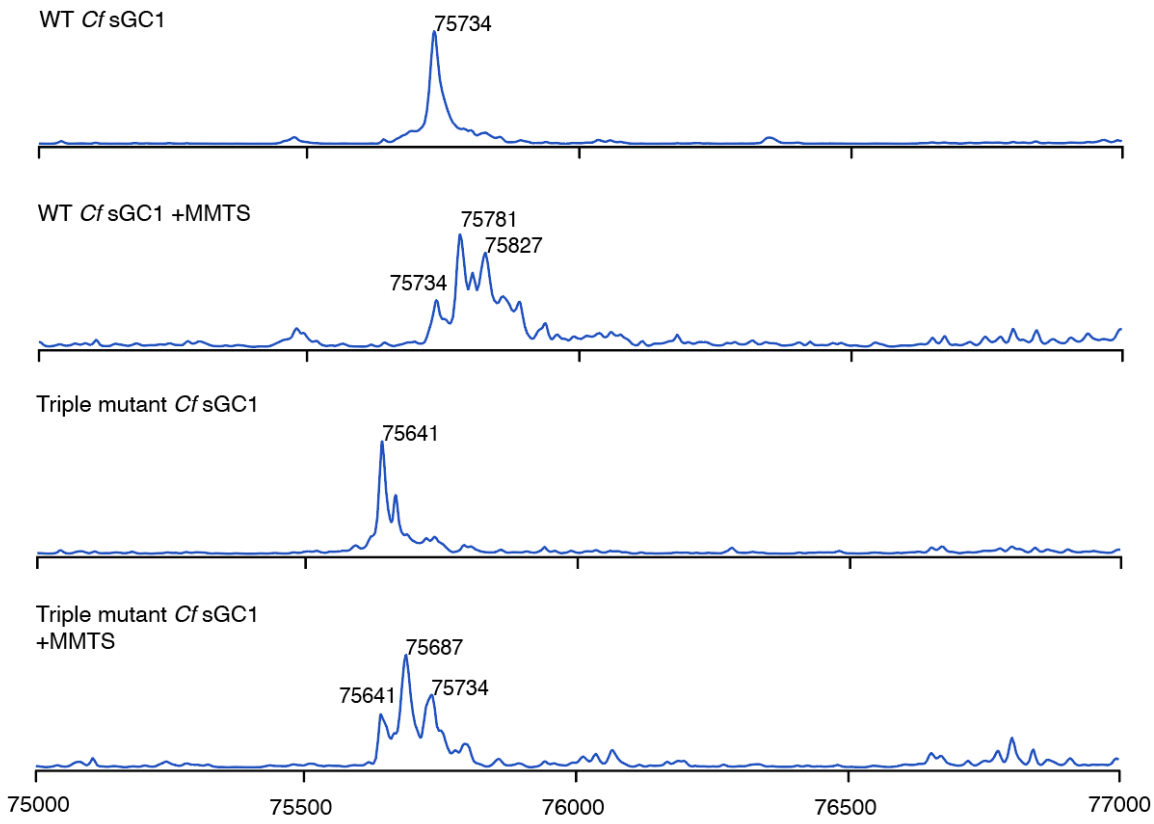


Figure 4.6. Mass spectrometry of MMTS labeled *Cf* sGC1. Expected molecular weight: WT *Cf* sGC1: 75732; Triple variant: 75638. MMTS labeling results in an S-methyl adduct; expected molecular weight of the adduct is +46 Da. MS results indicate that at 5 µM, both *Cf* sGC1 WT and triple variant proteins were primarily single-labeled with a small portion non labeled or double labeled. As a homodimer, the protein is primarily in the labeled state. MS results were visualized using Chartograph.

Conclusions

The non-heme interaction of sGC towards NO is a crucial mechanistic detail of sGC activation. Its importance could extend beyond the realm of sGCs, as a transient non-heme NO interaction may be utilized by many other proteins in diverse pathways to sense and respond to a NO signal. On the other hand, sGC is a well-studied drug target for diseases of the vasculature, so information on how sGC can be activated by excess NO will likely benefit development of novel and more effective therapeutics. High resolution structures of α/β-type sGC and access to *Cf* sGC1, a NO-activated sGC tractable to heterologous expression are invaluable tools that enabled long standing hypotheses of non-heme NO activation to be directly tested.

Our data cannot disprove that the cysteines at the three positions, 123, 355 and 489 are responsible for non-heme activation of *Cf* sGC1 but suggests that they are certainly not solely responsible. On the other hand, the three positions are likely not responsible for the MMTS-mediated inhibition of *Cf* sGC1. This is despite the fact that cysteines at equivalent positions are well conserved across α/β-type sGC and aligned well with the sequence of *Cf* sGC1. In future experiments, conservation of cysteines in similar structural regions of the sGC can be further

explored. This approach can be used to reveal residues like C177 that are in a region of the sequence that do not align well with the consensus sequence of α/β-type sGC but may play similar structural or biochemical roles as cysteines well conserved in α/β-type sGC. Additionally, given that *Cf* sGC1 is tractable to expression in *E. coli* and appears resilient to mutagenesis, constructs that employ more aggressive mutagenic strategies can be tested. One such construct is the cysteine-free variant of *Cf* sGC1. This approach can allow for the direct testing of the hypothesis that cysteines are responsible for the full activation of *Cf* sGC1.

References

(1) Wood, K. C.; Batchelor, A. M.; Bartus, K.; Harris, K. L.; Garthwaite, G.; Vernon, J.; Garthwaite, J. Picomolar Nitric Oxide Signals from Central Neurons Recorded Using Ultrasensitive Detector Cells. *J. Biol. Chem.* **2011**, *286* (50), 43172–43181

(2) Arancio, O.; Kiebler, M.; Lee, C. J.; Lev-Ram, V.; Tsien, R. Y.; Kandel, E. R.; Hawkins, R. D. Nitric Oxide Acts Directly in the Presynaptic Neuron to Produce Long-Term Potentiation in Cultured Hippocampal Neurons. *Cell* **1996**, *87* (6), 1025–1035

(3) Ignarro, L. J.; Byrns, R. E.; Buga, G. M.; Wood, K. S. Endothelium-Derived Relaxing Factor from Pulmonary Artery and Vein Possesses Pharmacologic and Chemical Properties Identical to Those of Nitric Oxide Radical. *Circ. Res.* **1987**, *61* (6), 866–879

(4) Kang, Y.; Liu, R.; Wu, J.-X.; Chen, L. Structural Insights into the Mechanism of Human Soluble Guanylate Cyclase. *Nature* **2019**, *574* (7777), 206–210

(5) Horst, B. G.; Yokom, A. L.; Rosenberg, D. J.; Morris, K. L.; Hammel, M.; Hurley, J. H.; Marletta, M. A. Allosteric Activation of the Nitric Oxide Receptor Soluble Guanylate Cyclase Mapped by Cryo-Electron Microscopy. *eLife* **2019**, 8:e50634

(6) Wittenborn, E. C.; Marletta, M. A. Structural Perspectives on the Mechanism of Soluble Guanylate Cyclase Activation. *Int. J. Mol. Sci.* **2021**, *22* (11), 5439

(7) Liu, R.; Kang, Y.; Chen, L. Activation Mechanism of Human Soluble Guanylate Cyclase by Stimulators and Activators. *Nat. Commun.* **2021**, *12* (1), 5492

(8) Cary, S. P. L.; Winger, J. A.; Marletta, M. A. Tonic and Acute Nitric Oxide Signaling through Soluble Guanylate Cyclase Is Mediated by Nonheme Nitric Oxide, ATP, and GTP. *Proc. Natl. Acad. Sci. U S A* **2005**, *102* (37), 13064–13069

(9) Russwurm, M.; Koesling, D. NO Activation of Guanylyl Cyclase. *EMBO J.* **2004**, *23* (22), 4443–4450

(10) Zhao, Y.; Brandish, P. E.; Ballou, D. P.; Marletta, M. A. A Molecular Basis for Nitric Oxide Sensing by Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U S A* **1999**, *96* (26), 14753–14758

(11) Guo, Y.; Suess, D. L. M.; Herzik, M. A.; Iavarone, A. T.; Britt, R. D.; Marletta, M. A. Regulation of Nitric Oxide Signaling by Formation of a Distal Receptor–Ligand Complex. *Nat. Chem. Biol.* **2017**, *13* (12), 1216–1221

(12) Bellamy, T. C.; Wood, J.; Goodwin, D. A.; Garthwaite, J. Rapid Desensitization of the Nitric Oxide Receptor, Soluble Guanylyl Cyclase, Underlies Diversity of Cellular CGMP Responses. *Proc. Natl. Acad. Sci. U S A* **2000**, *97* (6), 2928–2933

(13) Derbyshire, E. R.; Marletta, M. A. Butyl Isocyanide as a Probe of the Activation Mechanism of Soluble Guanylate Cyclase: Investigating the Role of Non-Heme Nitric Oxide. *J. Biol. Chem.* **2007**, *282* (49), 35741–35748

(14) Fernhoff, N. B.; Derbyshire, E. R.; Marletta, M. A. A Nitric Oxide/Cysteine Interaction Mediates the Activation of Soluble Guanylate Cyclase. *Proc. Natl. Acad. Sci. U S A* **2009**, *106* (51), 21602–21607

(15) Reyes-Rivera, J.; Wu, Y.; Guthrie, B. G. H.; Marletta, M. A.; King, N.; Brunet, T. Nitric Oxide Signaling Controls Collective Contractions in a Colonial Choanoflagellate. *Curr. Biol.* **2022**, *32* (11), 2539-2547.e5

(16) Horst, B. G.; Stewart, E. M.; Nazarian, A. A.; Marletta, M. A. Characterization of a Carbon Monoxide-Activated Soluble Guanylate Cyclase from *Chlamydomonas reinhardtii*. *Biochemistry* **2019**, *58* (17), 2250–2259

(17) Horst, B. G. The Structure and Activation of Soluble Guanylate Cyclase, University of California, Berkeley, 2019

(18) Atkinson, H. J.; Morris, J. H.; Ferrin, T. E.; Babbitt, P. C. Using Sequence Similarity Networks for Visualization of Relationships Across Diverse Protein Superfamilies. *PLoS One* **2009**, *4* (2), e4345

Chapter 5
Survey of O_2-sensing sGC of *C. flexa*

Abstract

Detection and response to O_2 levels in the environment is necessary for aerobic life. In eukaryotes, gas sensing is partially mediated by the hemoprotein soluble guanylate cyclase (sGC) that produces the secondary messenger cGMP. While some sGCs are highly specific sensors for NO, there are many sGCs that are predicted to bind O_2. The molecular details of ligand-mediated activity regulation in O_2-binding sGC are not well understood. The genome of the choanoflagellate *Choanoeca flexa* encodes four sGCs, and three are predicted by sequence alignment to bind O_2. Reported here is the characterization of *Cf* sGC4, an O_2-binding sGC from *C. flexa*. *Cf* sGC4 binds O_2 as well as NO and CO. All three ligands lead to a 6-coordiated Fe(II) heme complex. *Cf* sGC4 is active as a homodimer (566 ± 54 nmol/min/mg) and this activity is inhibited by ligand binding. Small angle X-ray scattering experiments indicate that *Cf* sGC4 can form an extended homodimer with R_g = 55.2 ± 0.3 Å. Conformational change was not observed in *Cf* sGC4 in the presence of NO. The ortholog of *Cf* sGC4 that was predicted to be NO specific from the sister species of *C. flexa*, *Choanoeca perplexa* was also characterized. *Cp* sGC1 binds NO and not O_2. NO binding did not increase the activity of *Cp* sGC1. These data suggest that *Cf* sGC4 may not rely on conformational extension to regulate activity, but additional experiments are necessary to understand the full mechanism of ligand effect on activity.

This work was done in collaboration with Dr. William C. Thomas from the lab of Prof. Michael A. Marletta, who contributed SAXS data analyses.

Introduction

Many organisms employ heme-based biosensors for gas molecules, including those that sense O_2, to detect and respond to changing environmental conditions. In eukaryotes, soluble guanylate cyclases (sGC) are gas sensing hemoproteins best known for sensing nitric oxide and synthesizing the second messenger molecule cyclic guanosine monophosphate (cGMP).[1] The best studied type of sGC, α/β-type sGC from mammals and insects, are highly selective sensors of nitric oxide and lack the ability to bind oxygen.[2] Selectivity against O_2 is a crucial feature of a NO sensing pathway, as O_2 is present in the cellular environment at a concentration several orders of magnitude over that of NO. However, studies involving bacterial homologs of the heme-nitric oxide-oxygen binding (H-NOX) domain of sGC identified standalone H-NOX domains with nanomolar binding affinity to O_2 and prompted studies to elucidate the mechanism for gas ligand differentiation.[3,4] The selectivity for NO and against O_2 using a hemoprotein would appear to be particularly difficult. The crystal structure of *Caldenaerobacter subterraneous* subsp. *tengcongensis* O_2-binding H-NOX protein pointed to a distal pocket Tyr residue as a critical residue for enabling O_2 binding through a specific and strong hydrogen bond.[5] Mutagenesis studies to incorporate a Tyr residue at the equivalent position in the NO-selective *Legionella pneumophila* H-NOX protein enabled it to bind O_2.[6] These observations established that a distal pocket Tyr residue is necessary for enabling O_2 binding in H-NOX protein and allowed for predictions of ligand preference of a sGC to be made based on sequence information.

Initial work carried out in model organisms *Caenorhabditis elegans, Manduca sexta* and *Drosophila melanogaster* led to the discovery and characterization of O_2-sensing sGCs and elucidated their connections to animal behavior.[7–10] Aerotaxis, physical migration towards preferred O_2 levels, is a behavior frequently found in bacteria and eukaryotes alike.[11] An example of aerotaxis in animals can be found in *C. elegans*, where the worms migrate to a preferred O_2 concentration of 7 - 12% to avoid hyperoxia.[7,12] In collaboration with the Bargmann lab, the Marletta lab reported that knockout of the neuron-expressed sGC GCY-35 resulted in the loss of avoidance behavior towards high O_2 concentration, suggesting that GCY-35 is a critical regulatory component of aerotaxis.[7] While the H-NOX domain truncation of GCY-35 was shown to bind O_2, biochemical evidence was lacking to link GCY-35 activity directly to O_2 avoidance. Discovery of O_2-binding sGCs in neuronal cells of *Drosophila melanogaster* led to the characterization of *Dm* Gyc-88E.[9,13] To date, *Dm* Gyc-88E has been the only O_2-binding sGC that has been purified and characterized. Huang and colleagues demonstrated that *Dm* Gyc-88E binds diatomic gas ligands NO, CO and O_2, is active as a homodimer, and is inhibited by gas ligand binding.[14] *In vivo* evidence suggests that *Dm* Gyc-88E expresses as a physiological homodimer in the epithelial cells of the *Drosophila* mid gut, and the enzyme is hypothesized to mediate O_2 sensing there.[9]

Previously, we reported a novel, cGMP-mediated phenotype discovered in a colonial choanoflagellate, *Choanoeca flexa*.[15] BLAST search revealed four sequences that were homologs of sGC and were named *Cf* sGC1 to 4. Through sequence alignment, *Cf* sGC2, *Cf* sGC3 and *Cf* sGC4 were all predicted to be oxygen sensing sGCs. Reported in this chapter is the characterization of the predicted O_2-binding sGC, *Cf* sGC4. Like *Dm* Gyc-88E, *Cf* sGC4 binds NO, CO and O_2 ligands, and its activity is inhibited by ligand binding. A small angle X-ray scattering study was conducted and revealed that *Cf* sGC4 can form an extended homodimer but did not undergo conformational change in the presence of NO. An ortholog of *Cf* sGC4 from the sister species *Choanoeca perplexa* that was predicted to bind NO specifically was also characterized.

Methods

Materials

Primers and gBlock DNA were purchased from Integrated DNA Technologies. Gibson Master Mix and *E. coli* competent cell stocks were purchased from the UC Berkeley QB3 MacroLab. Phusion Polymerase was purchased from New England Biolabs. Plasmid extraction and DNA purification kits were purchased from Qiagen. Terrific Broth media powder, kanamycin, benzamidine and 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) were purchased from Research Products International. Sodium dithionite ($Na_2S_2O_4$), DNAse I from bovine pancreas, sodium phosphate, sodium acetate, magnesium acetate and guanosine 5'-triphosphate sodium salt were purchased from Sigma Aldrich. Isopropyl β-D-1-thiogalactopyranoside (IPTG), 4-(2-hydroxyethyl)piperazine ethanesulfonic acid (HEPES), ammonium acetate and acetonitrile (HPLC grade) were purchased from Thermo Fisher Scientific. Lysozyme from chicken egg white and sodium chloride were purchased from VWR Scientific. Dithiothreitol (DTT) was purchased from BACHEM. Imidazole was purchased from Oakwood Chemical. Vivaspin spin concentrators were purchased from Sartorius. Hemin chloride and d-aminolevulinic acid (d-ALA) were purchased from Spectrum Chemical Company. proliNONOate was purchased from Cayman Chemical Company.

Cloning expression constructs

First-strand cDNA was used as the template for cloning *Cf* sGC4(1-631). Forward and reverse primers were designed against 5' and 3' ends of the target transcript (transcript name: TRINITY_DN11295_c0_g2_i3). Forward: TAAGAAGGAGATATACCATGTACGGTTTCGTC GCCG; reverse: AATGGTGATGATGGTGATGAACAGATGCGTCTGGTGGTGTC. Underlined portions anneal to the cDNA template. The PCR product was inserted into a pET28b vector using Gibson assembly, and the fully assembled plasmid was verified by sequencing (UC Berkeley sequencing facility). gBlock DNA with 5' and 3' end overhang was purchased for cloning *Cp* sGC1 and inserted into a pET28b vector using Gibson assembly directly.

Protein expression and purification

Expression of *Cf* sGC4 was carried out in *E. coli* RP523(DE3) cells for increased heme incorporation,[16] and expression of *Cp* sGC1 was carried out in *E. coli* BL21(DE3) cells. Prior to large scale expression of *Cf* sGC4, test expression of transformed *E. coli* RP523 colonies was carried out using 50 mL small scale culture. Overnight culture was prepared by inoculating one colony in 5 mL of Terrific Broth media supplemented with 50 µg/mL kanamycin and 20 µg/mL hemin. After overnight shaking at 37 °C, the overnight culture was diluted 1:200 into Terrific Broth (TB) media supplemented with 50 µg/mL kanamycin and 20 µg/mL hemin. The 50 mL-scale culture in TB media was incubated at 37 °C until cell density reached OD_{600} = 0.6 – 0.8 before protein production was induced by the addition of 500 µM IPTG. After induction, the culture was incubated at 18 °C for 16 – 20 hours. Protein production from each colony was estimated by lysing equal quantity of cells determined through OD matching of post induction culture, and analysis using SDS-PAGE gel. The colony with the best protein production was then selected for 1 L scale expression. Large culture expression followed the same protocol as above with minor modifications. After the overnight incubation at 18 °C, the cell suspension was centrifuged at 4000 x g at 4 °C. The cell paste was collected, flash frozen in liquid nitrogen, and transferred to -80 °C for storage. *Cp* sGC1 did not require an expression test, and the 1 L scale expression followed the same protocol as *Cf* sGC4 with minor differences. The overnight culture of *Cp* sGC1 was prepared in Luria Broth supplemented with 500 µM $FeCl_3$ and 50 µg/mL kanamycin. The 1 L scale culture of *Cp* sGC1 was supplemented with 500 µM $FeCl_3$ and 50 µg/mL kanamycin, and 1 mM d-ALA, a heme precursor, was supplemented before protein production was induced.

Protein purification was carried out at 4 °C unless otherwise specified. Cell pellets were thawed in an ice water bath and resuspended in ice cold Buffer A (50 mM sodium phosphate, 150 mM NaCl, 5 mM imidazole, 5 % glycerol, pH 8.0) supplemented with 50 mM benzamidine, 0.2 mM AEBSF, 0.25 mg/mL DNAse I, and 0.25 mg/mL lysozyme. The cell resuspension was lysed using a high-pressure homogenizer (Avestin Emulsiflex C5). The resulting cell lysate was clarified by centrifugation (36,000 x g, 55 min) and passed through a gravity column loaded with His60 Ni Superflow Resin equilibrated with Buffer A. The column was subsequently washed with 10 CV Buffer A, 10 CV Buffer B (50 mM sodium phosphate, 150 mM NaCl, 40 mM imidazole, 5 % glycerol, pH 8.0) and His-tagged protein was eluted with 5 CV Buffer C (50 mM sodium phosphate, 150 mM NaCl, 400 mM imidazole, 5 % glycerol, pH 8.0). Protein-containing fractions determined by SDS-PAGE electrophoresis were combined and concentrated using a 50 kDa molecular weight cut-off spin concentrator. Concentrated protein was loaded on a to a HiLoad 26/600 Superdex 200 size exclusion column (Cytiva) pre-equilibrated with Buffer D (50 mM Triethanolamine, 150 mM

NaCl, 5 mM DTT, 5% glycerol). Fractions with heme absorbance were concentrated to 25 – 50 μM and flash frozen in liquid nitrogen. Protein aliquots were stored at -80 °C for future use.

UV-vis spectroscopy

Cf sGC4 and *Cp* sGC1 were handled in an argon-filled glove bag. Protein-bound heme was reduced by adding $Na_2S_2O_4$ to a final concentration of 5 mM. Excess $Na_2S_2O_4$ was removed by gel filtration of the protein sample into Buffer E (50 mM HEPES, 150 mM NaCl, 5% glycerol, pH 7.5) using a pre-equilibrated Zeba spin desalting column (Thermo Scientific). A ferrous, ligand-free UV-vis absorption spectrum was collected on a Cary 300 UV-vis Spectrophotometer (Agilent Technologies). Fe(II)-NO bound protein was generated by adding the NO-releasing molecule proliNONOate (~10-fold excess) to the protein sample, Fe(II)-CO bound *Cf* sGC4 and *Cp* sGC1 were generated by adding CO-sparged Buffer E to the protein sample, and Fe(II)-O_2 bound *Cf* sGC4 and *Cp* sGC1 were generated by adding O_2-sparged buffer E to the protein sample before collecting the spectra.

Activity assay

Initial rates of *Cf* sGC4 and *Cp* sGC1 were measured by quantifying the production of cGMP in endpoint assays performed at 25 °C under anoxic conditions. Prior to the assay, protein aliquots were reduced using 10 mM $Na_2S_2O_4$ and buffer exchanged into Ar-sparged Buffer E (50 mM HEPES, 150 mM NaCl, 5 % glycerol, pH 7.5, 0.22 μm filtered) using a pre equilibrated 7.5 kDa MWCO Zeba Spin Desalting column (Thermo Scientific). Concentration of the protein was determined by the absorbance at 280 nm (*Cf* sGC4: $\varepsilon = 98{,}560\ M^{-1}\ cm^{-1}$; *Cp* sGC1: $\varepsilon = 83{,}660\ M^{-1}\ cm^{-1}$). Activity assay mixtures were prepared in Ar-sparged Buffer F supplemented with 5 mM DTT and 3 mM $MgCl_2$ and included 20 nM *Cf* sGC4 and 100 nM *Cp* sGC1 protein. For the NO-bound sample, NO was supplemented by the addition of 50 μM proliNONOate. For CO and O_2-bound samples, CO and O_2 were added through 1:5 dilution of CO and O_2-sparged buffer, respectively. Reaction was initiated by the addition of 1.5 mM GTP. For *Cf* sGC4, a 10-minute endpoint assay was performed. For *Cp* sGC1, a time course over 60 minutes was collected. The reaction was quenched by mixing 100 μL reaction volume with 10 μL 1% formic acid. The results of the enzymatic reaction were analyzed using reverse phase HPLC as previously reported.[17] Briefly, the reaction timepoint samples were injected onto an Eclipse Plus C18 column (4.6 x 100 mm, 3.5 μm particle size, Agilent). A gradient composed of 20 mM ammonium acetate, 0.1 % formic acid (Buffer F) and 99.9 % acetonitrile, 0.1 % formic acid (Buffer G) was used to separate GTP and cGMP. The gradient was as follows: 0 – 6 min, 2% Buffer G; 6 – 7.5 min, 25 – 100% Buffer G; 7.5 – 8.5 min, 100 % Buffer G. Concentrations of cGMP were determined from the peak area eluting at 4.2 minutes using a standard curve constructed from 1.25 – 20 μM cGMP standards.

Size exclusion chromatography-small angle X-ray scattering

Small-angle X-ray scattering (SAXS) was performed on *Cf* sGC4 under anaerobic conditions in a Coy glovebox at the MacChess BioSAXS beamline (Sector 7A) at the Cornell High-Energy Synchrotron Source, Ithaca, New York.[18–20] For both samples, *Cf* sGC4 was thawed and reduced directly in the glovebox, then exchanged into SAXS assay buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1% glycerol, 1 mM TCEP) using a pre-equilibrated Zeba spin desalting column (Thermo Scientific). Heme reduction was confirmed via in-line UV-vis spectroscopy (AvaSpec-ULS2048, Avantes), with the UV flow cell located directly after the SAXS sample cell. 100 μL of ~100 μM sample was injected onto a Superdex 200 10/300 column (Cytiva Life Sciences) pre-

equilibrated with SAXS assay buffer, and the column was run using an external pump at 0.5 mL/s (LC-20AD, Shimadzu). The xsNO sample was performed under similar conditions but with the addition of small molecule NO releasing agents, referred to as NONOates.[21] 100 μM proliNONOate was added directly to the sample prior to loading. Further, to ensure that the protein did not exchange into NO-free buffer on the column, the column was pre-equilibrated with A: 100 μM DETA NONOate (half-life 56 hours at 22 – 25 °C) B: 100 μM DEA NONOate (half-life 16 min at 22 – 25 °C) 40 minutes before loading.

Data were collected using a Dectris EIGER 4M detector at 4 °C. 1,500 × 2s X-ray exposures were collected for each run. Scattering images were integrated about the beam center and normalized by transmitted intensities measured on a photodiode beamstop. The X-ray wavelength of the experiment was $\lambda = 1.1061$ Å and the sample-to-detector distance was 1,748.0 mm, as determined by silver behenate calibration. Scattering was collected over a range of $q = 0.01$ Å^{-1} to 0.3 Å^{-1}, where q is the scattering vector $q = 4\pi\sin\theta/\lambda$ and 2θ is the scattering angle. Data processing, analysis, and comparison to predicted scattering profiles were performed in BioXTAS RAW using established protocol.[22,23] Background subtraction was performed by subtracting the buffer baseline prior to peak elution. Principal components of the elution were deconvolved using Evolving Factor Analysis (EFA) as implemented in BioXTAS RAW, which EFA enables mathematical separation of partially co-eluting species using iterative singular value decomposition to identify the numbers of distinguishable eluting species.[24] An initial component eluting prior to the main peak was identified as an aggregate with properties consistent with a filament. A second component was identified as minor, smaller aggregates but are still larger than a dimer. The main component was also computationally separated from a smaller species eluting in the shoulder of the main peak. Radii of gyration (R_g) were estimated with Guinier analysis. Error bars associated with R_g values are curve-fitting uncertainties from Guinier analysis. The calculated scattering of the predicted *Cf* sGC4 homodimer was determined with FoXS.[25,26] Low-resolution electron density maps were calculated using DENSS, as implemented in RAW.[27]

Predicted structure model

The sequence of *Cf* sGC4(1-631) and *Cp* sGC1(1-634) were used as input using the Google AlphaFold web server powered by AlphaFold 3.[28] Default settings for multimeric protein predictions were used. The rank order 1 structure was used to calculate predicted SAXS scattering, but no significant differences were observed between different rank order predictions.

Results and Discussion

Sequence analysis of O_2-sensing sGC from C. flexa

Four sequences from the transcriptome of *C. flexa* were annotated as sGCs. The sequence alignment of all four *C. flexa* sGC sequences was compared to a prototypical α/β-type sGC, *Rattus norvegicus* sGC, and the O_2-binding sGC previously characterized, *Drosophila melanogaster* Gyc-88E (Figure 5.1A, Figure 5.1B). *Cf* sGCs share the typical domain architecture of sGCs (Figure 5.1C). Like *Dm* Gyc-88E, all four *Cf* sGCs also have a C-terminal tail varying in length (Figure 5.1C). It is not known if this C-terminal tail exists in *Cf* sGCs *in vivo*, or what biochemical function they may serve.

For an sGC to be active as a homodimer, two sequence features must be present: 1. A binding site capable of binding heme in the H-NOX domain, and 2. Presence of the full set of residues needed for catalysis in the CAT domain. Sequence alignment indicates that *Cf* sGC2, 3 and 4 contain necessary residues for binding heme, and all three sGC have the full set of residues needed for catalysis (Figure 5.1A, Figure 5.1B). Therefore, we predicted that all three *Cf* sGCs listed above would be active as homodimers. Furthermore, *Cf* sGC2, 3 and 4 all harbor the tyrosine residue in the distal pocket of the H-NOX domain (Figure 5.1A). This tyrosine residue is absolutely conserved in O_2-binding H-NOX domains and is present in the O_2-binding sGC, *Dm* Gyc-88E. Given this, we predicted that *Cf* sGC2, 3 and 4 would all be capable of binding O_2.

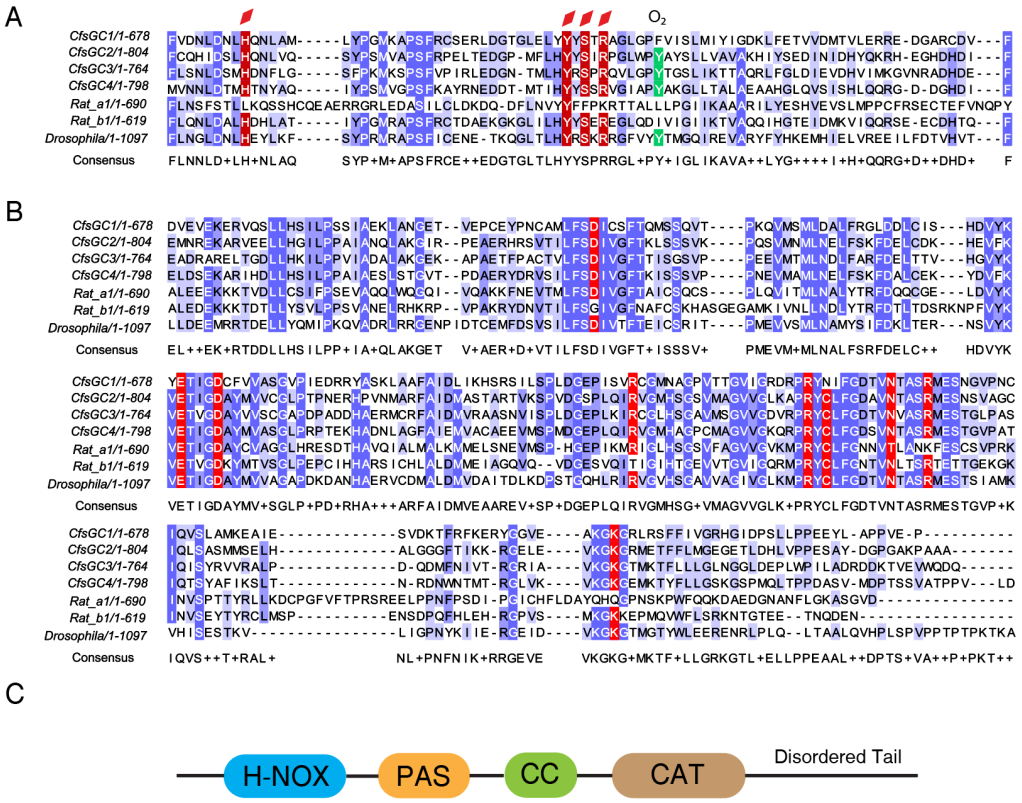


Figure 5.1. Sequence alignment and domain architecture of sGC from *C. flexa.* (A) H-NOX domain alignment suggests that all four *Cf* sGCs contain necessary residues (highlighted with the red diamond) for binding heme. Additionally, based on the presence of a conserved Tyr residue (green), *Cf* sGC2, 3, and 4 are predicted to bind O_2. (B) CAT domain alignment suggests that *Cf* sGCs contain necessary residues for catalysis. Therefore all four *Cf* sGCs are predicted to be active as homodimers. (C) *Cf* sGCs share the domain architecture of an α/β-type sGC. H-NOX, Heme-Nitric Oxide-Oxygen binding domain. PAS, Per-Arnt-Sim-like domain. CC, coiled coil domain. CAT, catalytic domain. Notably, like *Drosophila* Gyc-88E, all four *Cf* sGCs have a C-terminal disordered tail of unknown function.

Biochemical characterization of Cf sGC4

To obtain *Cf* sGC4 for biochemical characterization, a pET28b-based expression construct was prepared, and *Cf* sGC4 was expressed in *E. coli* BL21 cells initially. However, the protein was not expressed in the soluble fraction. We hypothesized that the C-terminal disordered region prevented proper folding of the protein during heterologous expression in *E. coli*. Based on the alignment with *Rn* sGC, a new expression construct was designed to truncate *Cf* sGC4 at position 631 (*Cf* sGC4(1-631)). Furthermore, to improve heme incorporation, the heme permeable *E. coli* strain RP523(DE3) was used.[16] The steps taken above allowed *Cf* sGC4(1-631) to express at a high yield in *E. coli* (Figure 5.2). The *Cf* sGC4(1-631) truncated variant was used in all subsequent biochemical experiments.

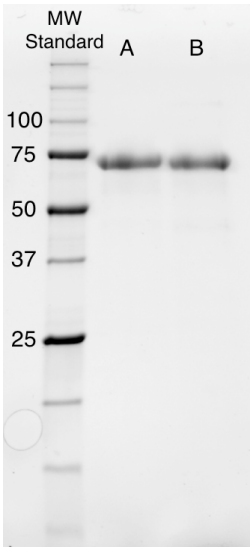


Figure 5.2. Post purification SDS-PAGE gel of *Cf* sGC4(1-631) and *Cp* sGC1(1-634). Lanes: A, *Cf* sGC4(1-631). B, *Cp* sGC1(1-634). The expected molecular weights of both proteins are around 70 kDa.

Cf sGC4 was expected to bind diatomic ligands NO and CO at its heme cofactor. Besides NO and CO, *Cf* sGC4 was also predicted to bind O_2. UV-vis absorption spectroscopy was used to investigate the ligand binding properties of *Cf* sGC4(1-631) (Figure 5.3A, Table 5.1). When *Cf* sGC4(1-631) was reduced under anaerobic conditions, it showed a Soret maximum of 426 nm, consistent with a 5-coordinate, unliganded heme. Unlike a NO-specific sGC, *Cf* sGC4(1-631) bound NO in a 6-coordinate state with a Soret maximum of 419 nm. *Cf* sGC4(1-631) also bound oxygen, showing a Soret maximum of 422 nm, consistent with a 6-coordinate Fe(II)-O_2 state. In addition to changes to the Soret band, the Q bands also showed increased separation during ligand binding for *Cf* sGC4(1-631). The UV-vis absorption wavelengths of *Cf* sGC4(1-631) and their comparisons to prototypical sGCs are summarized in Table 5.1.

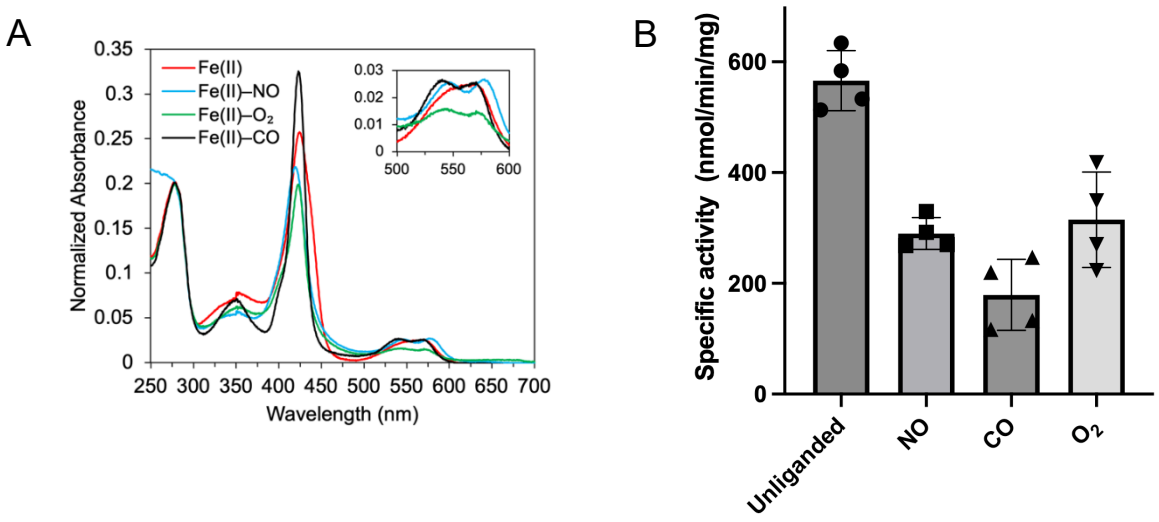


Figure 5.3. Gas ligand binding and ligand-bound activity of *Cf* sGC4(1-631). (A) UV-vis spectra of *Cf* sGC4(1-631) under unliganded and NO, CO and O_2-bound states. Soret maxima: Fe(II)-unliganded, 426 nm; Fe(II)-NO, 420 nm, Fe(II)-CO, 424 nm, Fe(II)-O_2, 423 nm. Inset, the Q band region showed increased separation upon ligand binding. (B) Activity of *Cf* sGC4(1-631). *Cf* sGC4(1-631) was inhibited by ligand binding. Data shown is the average of four replicates, and the error bar indicate one standard deviation.

Table 5.1. UV-vis absorption peaks of *Cf* sGC4 and *Cp* sGC1.

Ligand	Protein	Soret (nm)	β (nm)	α (nm)	Reference
None	α/β-type sGC	431	562		Stone, 1994[2]
	Cf sGC4 (1-631)	426	560		This work
	Cp sGC1 (1-634)	422	540	571	This work
	Dm Gyc-88E (1-597)	430	555		Huang, 2007[14]
NO	α/β-type sGC	399	537	572	Stone, 1994[2]
	Cf sGC4 (1-631)	420	547	580	This work
	Cp sGC1 (1-634)	417	548	578	This work
	Dm Gyc-88E (1-597)	419	544	572	Huang, 2007[14]
CO	α/β-type sGC	423	541	568	Stone, 1994[2]
	Cf sGC4 (1-631)	424	540	569	This work
	Cp sGC1 (1-634)	Not determined			
	Dm Gyc-88E (1-597)	421	543	570	Huang, 2007[14]
O_2	α/β-type sGC	No binding			Stone, 1994[2]
	Cf sGC4 (1-631)	423	545	572	This work
	Cp sGC1 (1-634)	No binding			This work
	Dm Gyc-88E (1-597)	416	556	591	Huang, 2007[14]

Both α/β-type sGC and *Cf* sGC1 sense NO specifically and are activated by NO binding. The O_2-binding sGC, *Dm* Gyc-88E, was previously found to be inhibited by ligand binding. When *Cf* sGC4(1-631) was assayed, like *Dm* Gyc-88E, was inhibited under the gas ligand-bound state (Figure 5.3B). The three ligands tested, NO, CO and O_2 inhibited *Cf* sGC4(1-631) between 2 to 3-fold. The distinctive ligand-responsive inhibition of *Cf* sGC4(1-631) and *Dm* Gyc-88E raises

intriguing questions about the conformational change that takes place after ligand binding, so a structural study of *Cf* sGC4(1-631) by small angle X-ray scattering (SAXS) was carried out.

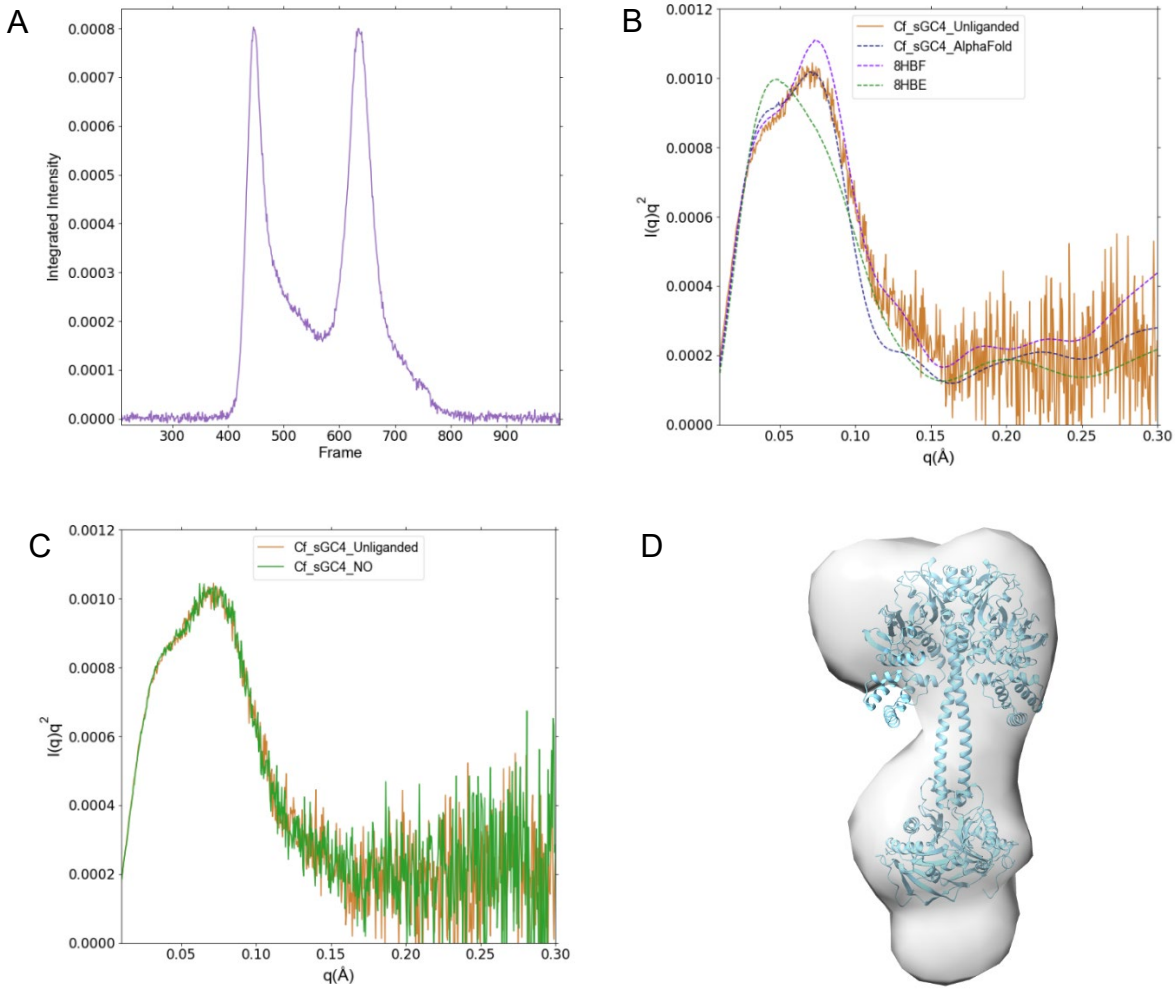


Figure 5.4. SAXS characterization of *Cf* sGC4(1-631). (A) SEC elution profile of *Cf* sGC4(1-631). The presence of a peak before the *Cf* sGC4(1-631) elution peak suggests the formation of large oligomers. (B) The experimentally determined SAXS profile of *Cf* sGC4(1-631) under NO-free condition was compared to the predicted SAXS profiles of the AlphaFold structure prediction of *Cf* sGC4(1-631), as well as that of human sGC unliganded (PDB: 8HBE) and active (PDB: 8HBF) structures. The SAXS profile of *Cf* sGC4(1-631) is a better fit to the predicted profile of the extended AlphaFold predicted structure, as well as the NO-induced extended human sGC active structure. (C) The SAXS profiles of unliganded and NO-bound *Cf* sGC4(1-631) were compared, and no significant difference was found between them. (D) The DENSS-calculated electronic density of *Cf* sGC4(1-631) was compared to the AlphaFold predicted structure of *Cf* sGC4(1-631). Overall, the AlphaFold structure fits well to the DENSS calculated density. The electron density near the C-terminal that is unaccounted for in the predicted structure can explain the discrepancy between the experimentally determined R_g to the calculated R_g.

SAXS study of Cf sGC4

As discussed in Chapter 1 and Chapter 3, for α/β-type sGC and *Cf* sGC1, NO binding triggers a large-scale conformational change that leads to an extension of the protein conformation. This conformational change can be studied using small angle X-ray scattering (SAXS). SAXS is a solution phase technique that provides information on the size and shape of the protein species

and is well suited to studying the conformational change that occurs in sGC activity regulation.[19,29,30] We expect that because *Cf* sGC4 is more active under the unliganded condition, it should be in a more extended conformation. Furthermore, we hypothesized that *Cf* sGC4 would undergo conformational change during inhibition and show a more compact conformation associated with the ligand bound, inhibited state. To test these hypotheses, SEC-SAXS study for *Cf* sGC4(1-631) was conducted under anaerobic conditions. The chromatogram of *Cf* sGC4(1-631) elution revealed two major elution peaks, indicating formation of at least two different oligomers of *Cf* sGC4(1-631) (Figure 5.4A). A linear algebra technique known as Evolving Factor Analysis (EFA) was used to computationally separate the components of the chromatogram (Figure C.1). The third component, mainly corresponding to the peak eluting at frame number 600 – 700, had a R_g within the range of an expected sGC dimer (R_g = 55.2 Å). We interpreted this peak to be *Cf* sGC4(1-631) in dimeric form. Both earlier components, mainly the peak between frames 400 – 500, were likely larger n-mers of *Cf* sGC4. The SAXS scattering profile of the third component was found to fit well to the predicted SAXS profile of the AlphaFold predicted structure of *Cf* sGC4(1-631) but did not fit well to either predicted scattering profiles of *Hs* sGC, under unliganded (PDB ID 8HBE) or NO-bound conditions (PDB ID 8HBF) (Figure 5.4B).[31] The predicted R_g of the AlphaFold predicted structure of *Cf* sGC4(1-631) was 46.5 Å. The discrepancy between the experimental and predicted R_g values may be caused by the EFA analysis not fully separating the dimeric component from the higher MW components.

Because *Cf* sGC4(1-631) can be inhibited by ligand binding, a second dataset was collected in which NO was added to the buffer. However, *Cf* sGC4(1-631) under the presence of NO did not show a conformational change based on SAXS analysis (Figure 5.4C). Furthermore, the elution profiles of the two samples were highly similar (Figure C.1A, Figure C.1E) and upon EFA, the two components were found to match in retention time (Figure C.1B, Figure C.1F). The affinity of *Cf* sGC4(1-631) for NO was not determined in this study, so two explanations are possible for this result: 1. *Cf* sGC4(1-631) did not undergo conformational change when NO was added, or 2. the NO concentration in our conditions was not sufficient to ensure 1:1 binding. Future SAXS experiments are necessary to definitively test the hypothesis regarding ligand-mediated conformational change in *Cf* sGC4.

Preliminary characterization of Cp sGC1

Choanoeca perplexa is a sister species of *C. flexa*.[32] Initially identified by Leadbeater in 1983, *C. perplexa* was also reported to form colonies, but the conditions that could reproducibly induce colony formation have not been identified.[33] A BLAST search of the transcriptome of *Choanoeca perplexa* was carried out and an ortholog of *Cf* sGC4 was identified to be named *Cp* sGC1.[34] Sequence alignment indicated that *Cp* sGC1 also has a C-terminal disordered tail, and the sequence identity to *Cf* sGC4 in the core sGC domains is 68.9 %. In comparison, *Cf* sGC4 shares between 30 and 40 % sequence identity with *Cf* sGC1, 2 and 3. Interestingly, despite the high sequence similarity, *Cp* sGC1 has a phenylalanine residue at the position of the O_2 binding-enabling tyrosine residue of *Cf* sGC4 and, therefore, was predicted to not bind O_2 (Figure 5.5A, Figure C.2A). Additionally, the CAT domain of *Cp* sGC1 contains all necessary residues for catalyzing the conversion of GTP to cGMP, suggesting that could be active as a homodimer (Figure C.2B). Since NO selectivity is often associated with ligand-dependent activation, if *Cp* sGC1 is also activated by NO, it may help pinpoint the region of the protein that determines whether sGC is activated or inhibited by ligand binding based on homology to *Cf* sGC4. Hence, *Cp* sGC1 was characterized *in vitro*. *Cp* sGC1 was expressed as a truncation containing residues 1-634 to remove

the C-terminal disordered tail for ease of heterologous expression. Here, gas ligand binding and ligand-bound activity of *Cp* sGC1(1-634) are reported.

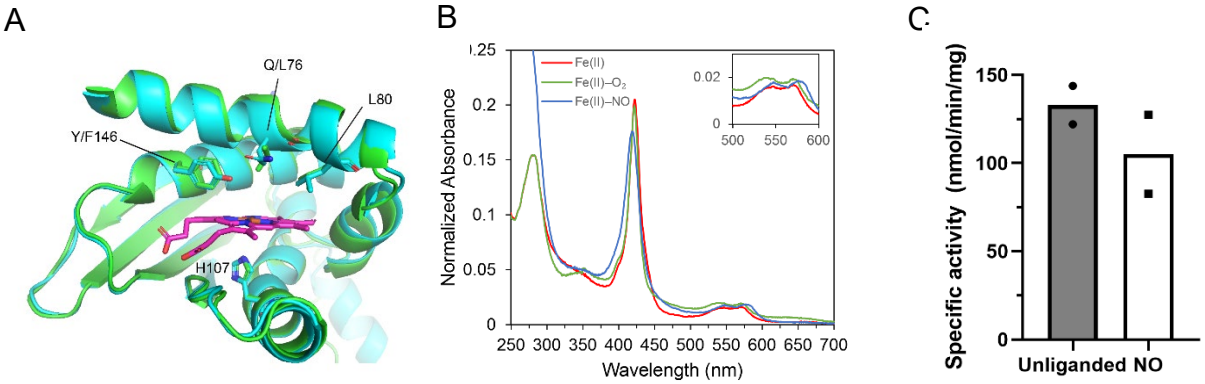


Figure 5.5. Preliminary characterization of *Cp* sGC1. (A) Heme binding pocket comparison between *Cp* sGC1 and *Cf* sGC4. AlphaFold predicted structures of the H-NOX domains of *Cf* sGC4 (cyan) and *Cp* sGC1 (Green) were aligned. Heme cofactor location was shown by alignment to human sGC, PDB 6JT2 and shown in purple. *Cp* sGC1 lacks essential hydrogen bonding residues for stabilizing Fe(II)-O_2 complex (shown in sticks) at positions Y146 and Q76 in *Cf* sGC4, where they are occupied by F146 and L76 in *Cp* sGC1. Both proteins have L80. (B) Gas ligand binding of *Cp* sGC1(1-634). *Cp* sGC1(1-634) binds NO to form a 6-coordinate Fe(II)-NO complex. No significant spectral change was seen when O_2 was added to a sample of reduced *Cp* sGC1(1-634). (C) Activity of *Cp* sGC1(1-634) with and without NO. *Cp* sGC1 isn't activated by NO binding.

UV-vis absorption spectroscopy was used to characterize the ligand-binding properties of *Cp* sGC1(1-634) (Figure 5.5B). Reduced *Cp* sGC1(1-634) showed a Soret maximum of 422 nm. This is noticeably different from a typical H-NOX domain, where the expected the Soret maximum is around 430 nm. Furthermore, the Q band region showed two peaks instead of one. However, when *Cp* sGC1(1-634) was treated with O_2, the UV-vis absorption spectrum did not noticeably shift. This can be explained by two possibilities: 1, *Cp* sGC1(1-634) binds O_2 so tightly that the treatment with dithionite cannot adequately remove the bound O_2 and reduce it to the Fe(II) unliganded state, or 2, *Cp* sGC1(1-634) does not bind O_2, and the unusual UV-vis absorption spectrum is a result of local electronic differences. The AlphaFold predicted structures of the H-NOX domains of *Cp* sGC1 and *Cf* sGC4 were then compared (Figure 5.5A). In *Cp* sGC1, residues required residues to stabilize Fe(II) heme-O_2 complex were not present. F146 and L76 were present at those key positions. Based on this, the second explanation stated above is more likely, but more experiments are needed to convincingly reach a conclusion. When reduced *Cp* sGC1(1-634) was treated with NO, the Soret maximum shifted to 419 nm, and a noticeable shift could also be seen in the Q band region. This suggests that when *Cp* sGC1(1-634) binds NO, a 6-coordinate, Fe(II) heme-NO complex forms. NO binding of *Cp* sGC1(1-634) is more akin to *Cf* sGC4 than *Cf* sGC1 and suggests that additional factors are involved in the strength of the proximal Fe-His bond and the effectiveness of the NO *trans* effect.

When the steady state activity of *Cp* sGC1(1-634) was assayed, NO binding did not significantly perturb its activity. Presence of O_2 also did not significantly affect the activity of *Cp* sGC1(1-634). Overall, the case of *Cp* sGC1 is a perplexing one, and it again raises questions about residues that control the *trans* effect of the NO ligand as well as the strength of the Fe-His bond.

It also raises questions about ligand activation or ligand inhibition. Although it is predicted by sequence to be active as a homodimer, it is possible that *Cp* sGC1 may interact with a partner subunit to form heterodimers, and this interaction may be the key to enabling ligand control of activity. The biologically relevant state of *Cp* sGC1 and its physiological role remain to be determined.

Conclusions

Taken together, these data show that *Cf* sGC4 is regulated by O_2 binding. Like *Dm* Gyc-88E, *Cf* sGC4 is active as a homodimer and is inhibited by gas ligand binding. Preliminary SAXS data suggests that the truncated *Cf* sGC4(1-631) forms a highly extended homodimer under unliganded condition as expected. However, preliminary SAXS data also show that *Cf* sGC4(1-631) did not undergo conformational change with NO bound. Whether the regulation of *Cf* sGC4 is related to large-scale conformational changes as seen in animal sGCs remains to be elucidated.

Further study on the regulation of ligand-induced activity change in O_2-binding sGC should yield interesting insight to the regulation of sGC activity. *Cf* sGC4 is an attractive new model system for studying conformational changes in O_2-binding sGC thanks to its tractability to heterologous expression. To address the issue of large oligomers forming, C-terminal truncations of *Cf* sGC4 can be screened for an active, non-oligomer forming variant and used in structural studies such as SAXS and cryo-electron microscopy. Additionally, sGCs that are predicted by sequence to bind O_2 are numerous, and this sequence space is ripe for exploration. It remains to be shown whether there exists in nature a class of O_2-binding sGC that can be activated by ligand binding.

Since choanoflagellates encode O_2-binding sGC, there could exist O_2-sensitive phenotypes mediated by sGC-produced cGMP. Notably, a choanoflagellate *Salpingoeca rosetta* has been discovered to exhibit aerotaxis behavior.[35] *S. rosetta* aerotaxis happens through stochastic navigation consistent with the "run-and-tumble" chemotaxis model, where control of the tumbling rate of *S. rosetta* colonies can bias the direction of movement towards a chemical stimulant, in this case, O_2.[35] The genome of *S. rosetta* encodes at least one O_2-sensing sGC, which could play a role in regulating this process. Screening for O_2 sensitive phenotypes specifically involving flagella and collars of *S. rosetta* and using a cell-permeable cGMP analog, 8-Br-cGMP to trigger these phenotypes can be the first steps to probe the connection between O_2 sensing behavior of *S. rosetta* and sGC-cGMP signaling. In parallel, O_2-sensing sGC of *S. rosetta* can be expressed and characterized. *S. rosetta* sGCs alone may hold interesting information on the biochemistry and allosteric regulation of O_2-sensing sGCs.

References

(1) Stone, J. R.; Marletta, M. A. Spectral and Kinetic Studies on the Activation of Soluble Guanylate Cyclase by Nitric Oxide. *Biochemistry* **1996**, *35* (4), 1093–1099

(2) Stone, J. R.; Marletta, M. A. Soluble Guanylate Cyclase from Bovine Lung: Activation with Nitric Oxide and Carbon Monoxide and Spectral Characterization of the Ferrous and Ferric States. *Biochemistry* **1994**, *33* (18), 5636–5640

(3) Karow, D. S.; Pan, D.; Tran, R.; Pellicena, P.; Presley, A.; Mathies, R. A.; Marletta, M. A. Spectroscopic Characterization of the Soluble Guanylate Cyclase-like Heme Domains from *Vibrio Cholerae* and *Thermoanaerobacter tengcongensis*. *Biochemistry* **2004**, *43* (31), 10203–10211

(4) Nioche, P.; Berka, V.; Vipond, J.; Minton, N.; Tsai, A. L.; Raman, C. S. Femtomolar Sensitivity of a NO Sensor from *Clostridium botulinum*. *Science* **2004**, *306* (5701), 1550–1553

(5) Pellicena, P.; Karow, D. S.; Boon, E. M.; Marletta, M. A.; Kuriyan, J. Crystal Structure of an Oxygen-Binding Heme Domain Related to Soluble Guanylate Cyclases. *Proc. Natl. Acad. Sci. U S A* **2004**, *101* (35), 12854–12859

(6) Boon, E. M.; Huang, S. H.; Marletta, M. A. A Molecular Basis for NO Selectivity in Soluble Guanylate Cyclase. *Nat. Chem. Biol.* **2005**, *1* (1), 53–59

(7) Gray, J. M.; Karow, D. S.; Lu, H.; Chang, A. J.; Chang, J. S.; Ellis, R. E.; Marletta, M. A.; Bargmann, C. I. Oxygen Sensation and Social Feeding Mediated by a *C. elegans* Guanylate Cyclase Homologue. *Nature* **2004**, *430* (6997), 317–322

(8) Cheung, B. H. H.; Arellano-Carbajal, F.; Rybicki, I.; de Bono, M. Soluble Guanylate Cyclases Act in Neurons Exposed to the Body Fluid to Promote *C. elegans* Aggregation Behavior. *Curr. Biol.* **2004**, *14* (12), 1105–1111

(9) Morton, D. B. Atypical Soluble Guanylyl Cyclases in Drosophila Can Function as Molecular Oxygen Sensors. *J. Biol. Chem.* **2004**, *279* (49), 50651–50653

(10) Morton, D. B.; Anderson, E. J. MsGC-B3 Forms Active Homodimers and Inactive Heterodimers with NO-Sensitive Soluble Guanylyl Cyclase Subunits. *J. Exp. Biol.* **2003**, *206* (6), 937–947

(11) Taylor, B. L.; Zhulin, I. B.; Johnson, M. S. Aerotaxis and Other Energy-Sensing Behavior in Bacteria. *Annu. Rev. Microbiol.* **1999**, *53* (1), 103–128

(12) Zimmer, M.; Gray, J. M.; Pokala, N.; Chang, A. J.; Karow, D. S.; Marletta, Michael. A.; Hudson, M. L.; Morton, D. B.; Chronis, N.; Bargmann, C. I. Neurons Detect Increases and Decreases in Oxygen Levels Using Distinct Guanylate Cyclases. *Neuron* **2009**, *61* (6), 865–879

(13) Vermehren-Schmaedick, A.; Ainsley, J. A.; Johnson, W. A.; Davies, S.-A.; Morton, D. B. Behavioral Responses to Hypoxia in Drosophila Larvae Are Mediated by Atypical Soluble Guanylyl Cyclases. *Genetics* **2010**, *186* (1), 183–196

(14) Huang, S. H.; Rio, D. C.; Marletta, M. A. Ligand Binding and Inhibition of an Oxygen-Sensitive Soluble Guanylate Cyclase, Gyc-88E, from *Drosophila*. *Biochemistry* **2007**, *46* (51), 15115–15122

(15) Reyes-Rivera, J.; Wu, Y.; Guthrie, B. G. H.; Marletta, M. A.; King, N.; Brunet, T. Nitric Oxide Signaling Controls Collective Contractions in a Colonial Choanoflagellate. *Curr. Biol.* **2022**, *32* (11), 2539-2547

(16) Woodward, J. J.; Martin, N. I.; Marletta, M. A. An *Escherichia Coli* Expression–Based Method for Heme Substitution. *Nat. Methods* **2007**, *4* (1), 43–45

(17) Horst, B. G.; Stewart, E. M.; Nazarian, A. A.; Marletta, M. A. Characterization of a Carbon Monoxide-Activated Soluble Guanylate Cyclase from *Chlamydomonas reinhardtii*. *Biochemistry* **2019**, *58* (17), 2250–2259

(18) Acerbo, A. S.; Cook, M. J.; Gillilan, R. E. Upgrade of MacCHESS Facility for X-Ray Scattering of Biological Macromolecules in Solution. *J. Synchrotron Radiat.* **2015**, *22* (1), 180–186

(19) Skou, S.; Gillilan, R. E.; Ando, N. Synchrotron-Based Small-Angle X-Ray Scattering of Proteins in Solution. *Nat. Protoc.* **2014**, *9* (7), 1727–1739

(20) Illava, G.; Gillilan, R.; Ando, N. Development of In-Line Anoxic Small-Angle X-Ray Scattering and Structural Characterization of an Oxygen-Sensing Transcriptional Regulator. *J. Biol. Chem.* **2023**, *299* (8), 105039

(21) Maragos, C. M.; Morley, D.; Wink, D. A.; Dunams, T. M.; Saavedra, J. E.; Hoffman, A.; Bove, A. A.; Isaac, L.; Hrabie, J. A.; Keefer, L. K. Complexes of · NO with Nucleophiles as Agents for the Controlled Biological Release of Nitric Oxide. Vasorelaxant Effects. *J. Med. Chem.* **1991**, *34* (11), 3242–3247

(22) Hopkins, J. B. *BioXTAS RAW 2* : New Developments for a Free Open-Source Program for Small-Angle Scattering Data Reduction and Analysis. *J. Appl. Crystallogr.* **2024**, *57* (1), 194–208

(23) Hopkins, J. B.; Gillilan, R. E.; Skou, S. *BioXTAS RAW* : Improvements to a Free Open-Source Program for Small-Angle X-Ray Scattering Data Reduction and Analysis. *J. Appl. Crystallogr.* **2017**, *50* (5), 1545–1553

(24) Meisburger, S. P.; Taylor, A. B.; Khan, C. A.; Zhang, S.; Fitzpatrick, P. F.; Ando, N. Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-Ray Scattering. *J. Am. Chem. Soc.* **2016**, *138* (20), 6506–6516

(25) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. FoXS, FoXSDock and MultiFoXS: Single-State and Multi-State Structural Modeling of Proteins and Their Complexes Based on SAXS Profiles. *Nucleic Acids Res.* **2016**, *44* (W1), W424–W429

(26) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. Accurate SAXS Profile Computation and Its Assessment by Contrast Variation Experiments. *Biophys. J.* **2013**, *105* (4), 962–974

(27) Grant, T. D. Ab Initio Electron Density Determination Directly from Solution Scattering Data. *Nat. Methods* **2018**, *15* (3), 191–193

(28) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J.; Bodenstein, S. W.; Evans, D. A.; Hung, C.-C.; O'Neill, M.; Reiman, D.; Tunyasuvunakool, K.; Wu, Z.; Žemgulytė, A.; Arvaniti, E.; Beattie, C.; Bertolli, O.; Bridgland, A.; Cherepanov, A.; Congreve, M.; Cowen-Rivers, A. I.; Cowie, A.; Figurnov, M.; Fuchs, F. B.; Gladman, H.; Jain, R.; Khan, Y. A.; Low, C. M. R.; Perlin, K.; Potapenko, A.; Savy, P.; Singh, S.; Stecula, A.; Thillaisundaram, A.; Tong, C.; Yakneen, S.; Zhong, E. D.; Zielinski, M.; Žídek, A.; Bapst, V.; Kohli, P.; Jaderberg, M.;

Hassabis, D.; Jumper, J. M. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, *630* (8016), 493–500

(29) Horst, B. G.; Yokom, A. L.; Rosenberg, D. J.; Morris, K. L.; Hammel, M.; Hurley, J. H.; Marletta, M. A. Allosteric Activation of the Nitric Oxide Receptor Soluble Guanylate Cyclase Mapped by Cryo-Electron Microscopy. *eLife* **2019**, *8*

(30) Wittenborn, E. C.; Thomas, W. C.; Houghton, K. A.; Wirachman, E. S.; Wu, Y.; Marletta, M. A. Role of the Coiled-Coil Domain in Allosteric Activity Regulation in Soluble Guanylate Cyclase. *Biochemistry* **2023**, *62* (10), 1568–1576

(31) Liu, R.; Kang, Y.; Chen, L. NO Binds to the Distal Site of Haem in the Fully Activated Soluble Guanylate Cyclase. *Nitric Oxide* **2023**, *134–135*, 17–22

(32) Brunet, T.; Larson, B. T.; Linden, T. A.; Vermeij, M. J. A.; McDonald, K.; King, N. Light-Regulated Collective Contractility in a Multicellular Choanoflagellate. *Science* **2019**, *366* (6463), 326–334

(33) Leadbeater, B. S. C. Life-History and Ultrastructure of a New Marine Species of *Proterospongia* (Choanoflagellida). *Journal of the Marine Biological Association of the United Kingdom* **1983**, *63* (1), 135–160

(34) Richter, D. J.; Fozouni, P.; Eisen, M. B.; King, N. Gene Family Innovation, Conservation and Loss on the Animal Stem Lineage. *eLife* **2018**, *7:e34226*

(35) Kirkegaard, J. B.; Bouillant, A.; Marron, A. O.; Leptos, K. C.; Goldstein, R. E. Aerotaxis in the Closest Relatives of Animals. *eLife* **2016**, *5:e18109*

Appendices

Appendix A

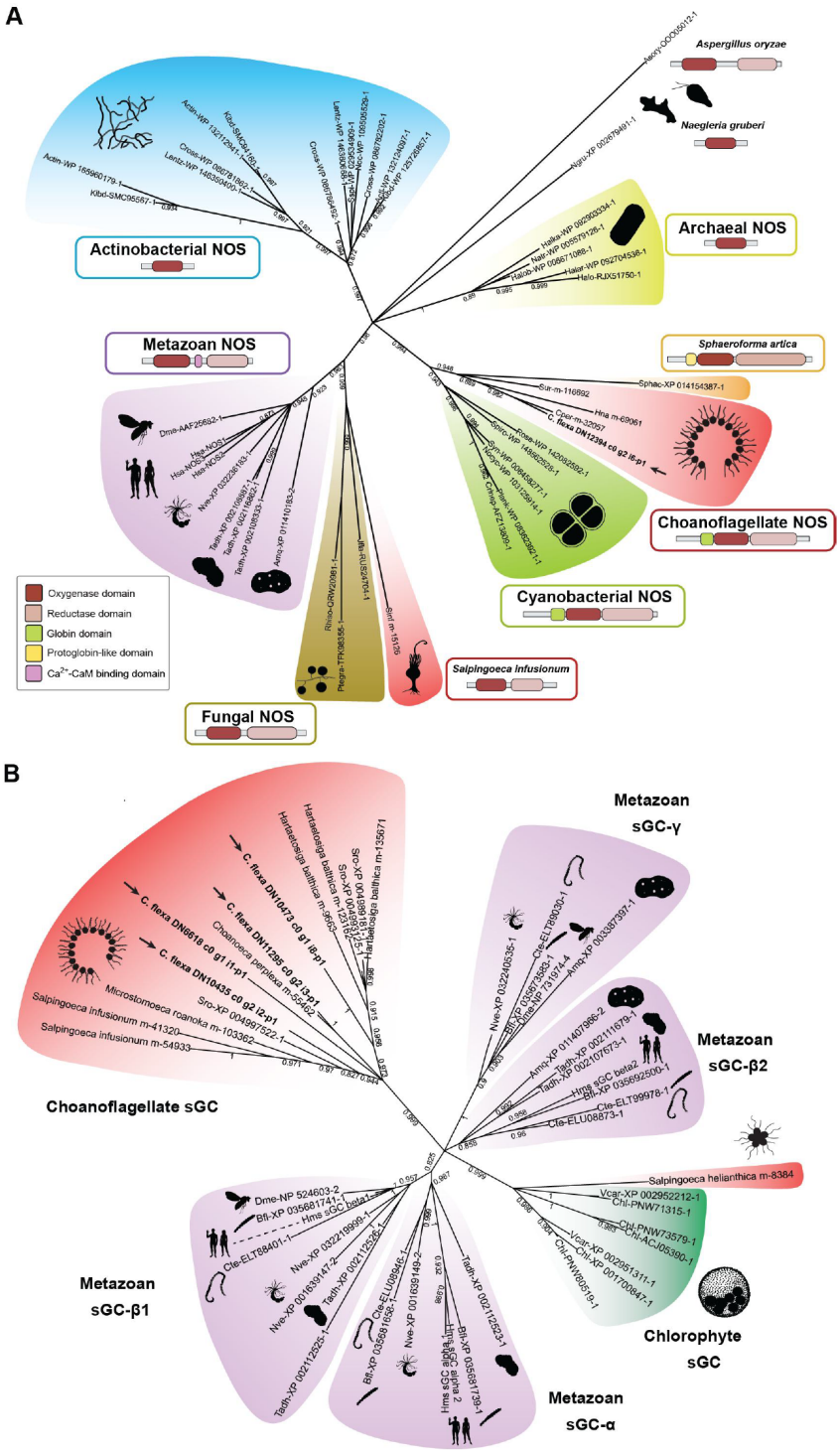


Figure A.1. Phylogenetic tree of NOS and sGC, related to Figure 2.2. (A) Phylogenetic tree of NOS from a selection of fully sequenced eukaryotic, archaeal, and bacterial genomes. Alignments and phylogenetic reconstruction were performed on the oxygenase domain only. The clades recovered tend to share common domain architectures out of the oxygenase domain (such as the reductase domain in eukaryotic NOSs and the globin domain in choanoflagellate and cyanobacterial NOSs), providing independent support to the phylogeny. The sister-group relationship between choanoflagellate and cyanobacterial NOSs as well as the shared domain architecture suggest a history of horizontal gene transfer. See Material and Methods for species name abbreviations. (B) Phylogenetic tree of sGC from a selection of fully sequenced animal and chlorophyte genomes, and choanoflagellate genomes and transcriptomes. Analysis was performed based on the full protein sequence. See Methods for species name abbreviations.

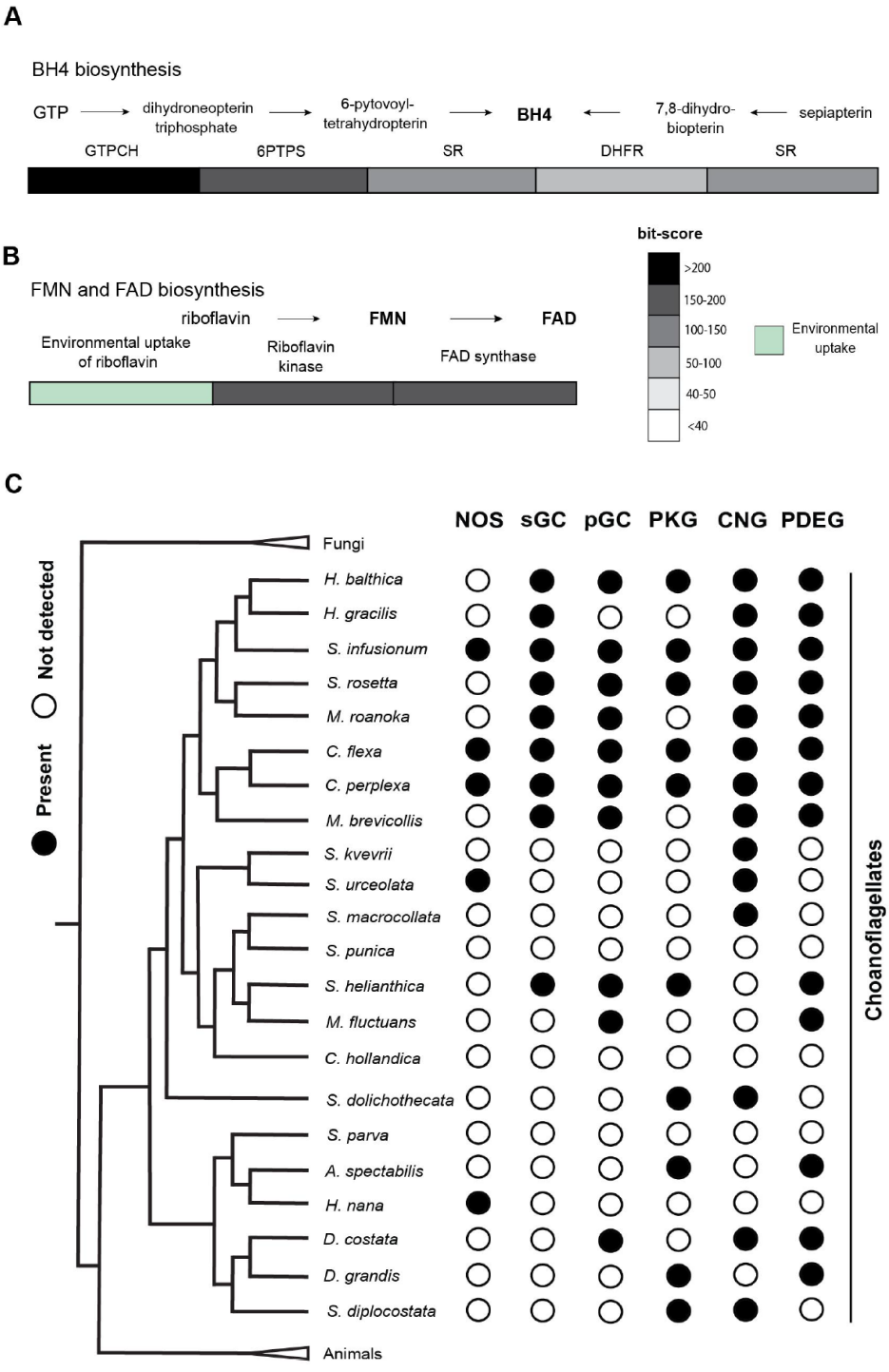

A
BH4 biosynthesis
GTP
dihydroneopterin triphosphate
6-pytovoyl-tetrahydropterin
BH4
7,8-dihydro-biopterin
sepiapterin
GTPCH
6PTPS
SR
DHFR
SR
B
FMN and FAD biosynthesis
riboflavin
FMN
FAD
Environmental uptake of riboflavin
Riboflavin kinase
FAD synthase
bit-score
>200
150-200
100-150
50-100
40-50
<40
Environmental uptake
C
NOS
sGC
pGC
PKG
CNG
PDEG
Not detected
Present
Fungi
H. balthica
H. gracilis
S. infusionum
S. rosetta
M. roanoka
C. flexa
C. perplexa
M. brevicollis
S. kvevrii
S. urceolata
S. macrocollata
S. punica
S. helianthica
M. fluctuans
C. hollandica
S. dolichothecata
S. parva
A. spectabilis
H. nana
D. costata
D. grandis
S. diplocostata
Animals
Choanoflagellates
D
120 - 178
C. flexa sGC1 SFRCSERLDGTGLELYYYSTRAGLGPFVISLMIYIGDKLFETVVDMTVLERRE-DGARCD
C. flexa sGC2 SFRPELTEDGP-MFLHYYSIRPGLWPYAYSLLVAVAKHIYSEDINIDHYQKRH-EGHDHD
C. flexa sGC3 SFVPIRLEDGN-TMLHYRSPRQVLGPYTGSLIKTTAQRLFGLDIEVDHVIMKGVNRADHD
C. flexa sGC4 SFKAYRNEDDT-MTIHYYSSRVGIAPYAKGLLTALAEAAHGLQVSISHLQQRG-D-DGHD
Human sGC SFRCTDAEKGKGLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQQRNE-ECDHT
Drosophila sGC SFICENE-TKQGLTLHYRSKRRGFVYYTMGQIREVARYFYHKEMHIELVREEILFDTVHV

Figure A.2. *C. flexa* encodes the complete biosynthetic pathways for the NOS cofactors and downstream NO/cGMP signaling components, related to Figure 2.2. **(A-B)** The *C. flexa* transcriptome was searched for genes encoding enzymes in the tetrahydrobiopterin (BH4; A) and FMN/FAD (B) biosynthesis pathways using BLASTP. For each step in the pathway, multiple bacterial, plant, fungal and/or animal genes were used as queries (see Methods), and the highest returned bit-score is shown. *C. flexa* encodes the complete biosynthetic pathway. Riboflavin is assumed to be uptaken from the environment.[1] (C) Phylogenetic distribution of guanylate cyclases (soluble: sGC, and membrane-bound: pGC), and cGMP signaling downstream components. In animals, cGMP is known to signal through cGMP-dependent kinases (PKG), cGMP-gated ion channels (CNG) and cGMP-dependent phosphodiesterases (PDEG). The three choanoflagellate species detected to express a NOS and sGC (*C. flexa*, *C. perplexa* and *S. infusionum*) also express PKG, CNG and PDEG. (D) Preferential binding predictions were made for *C. flexa* sGCs based on NO-sensitive and O_2- binding metazoan sGCs.[2–4] *C. flexa* sGC partial alignment with human sGC (NO-sensitive) and *Drosophila* sGC (O_2-binding). Amino acids important for heme binding are highlighted in grey. A distal tyrosine residue is highlighted in yellow, previously shown to be highly important for O_2 binding.[5] One NO-sensitive (sGC1) and three O_2-binding sGCs (sGC2, sGC3, sGC4) were predicted.

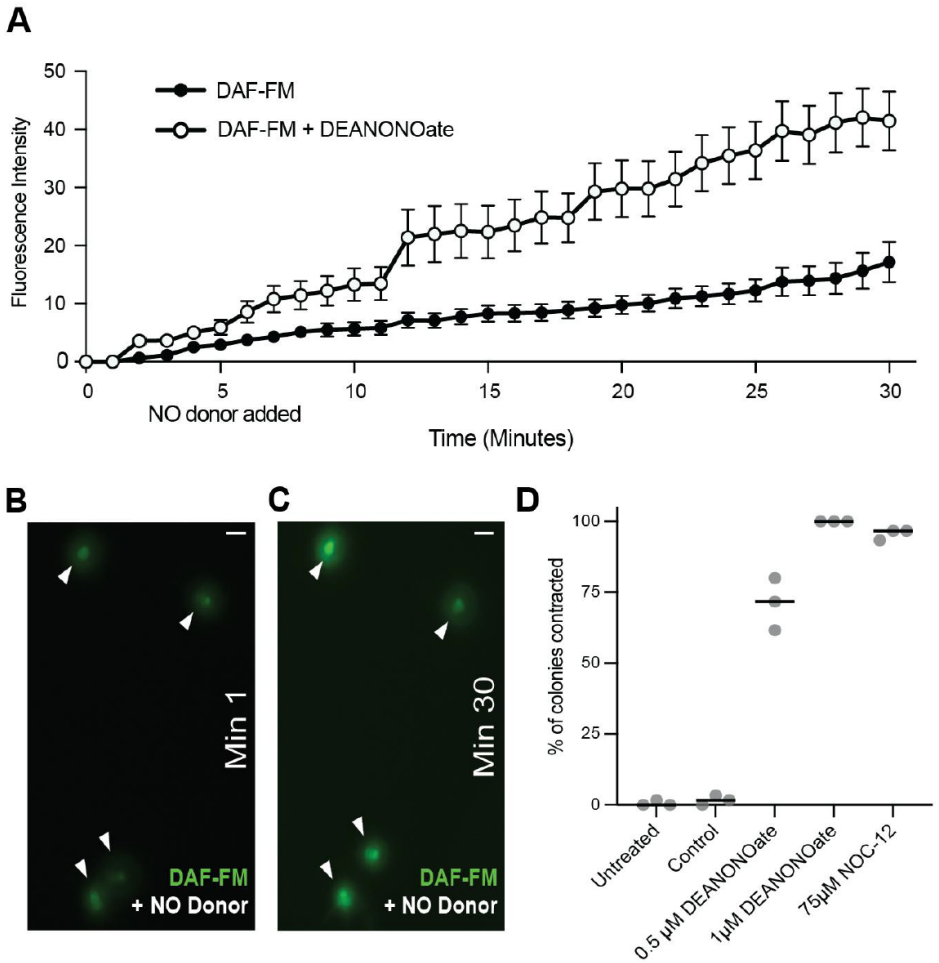


Figure A.3. Treatment with NO donors increases intracellular NO levels and induces colony contraction, related to Figure 2.3. (A-C) Intracellular NO levels over time. NO was labeled with the fluorescent probe DAF-FM and the fluoresce intensity was measured every minute for 30 minutes. NO donor DEA-NONOate was added at minute five. Intracellular fluorescence in NO donor-treated cells display a higher increase over time compared to control. Small increase in fluorescence in control group may be explained by unwashed dye incorporating inside of the cells and/or basal physiological NO levels. Intracellular intensity was measured using Image J software. (B-C) Representative micrographs are shown on the right. White arrows are pointing individual cells. Scale bar: 5 µM. NO donor was added at minute five. (D) NO donors DEA-NONOate and NOC-12 also induce colony contraction. Colonies from control group were treated with DMSO.

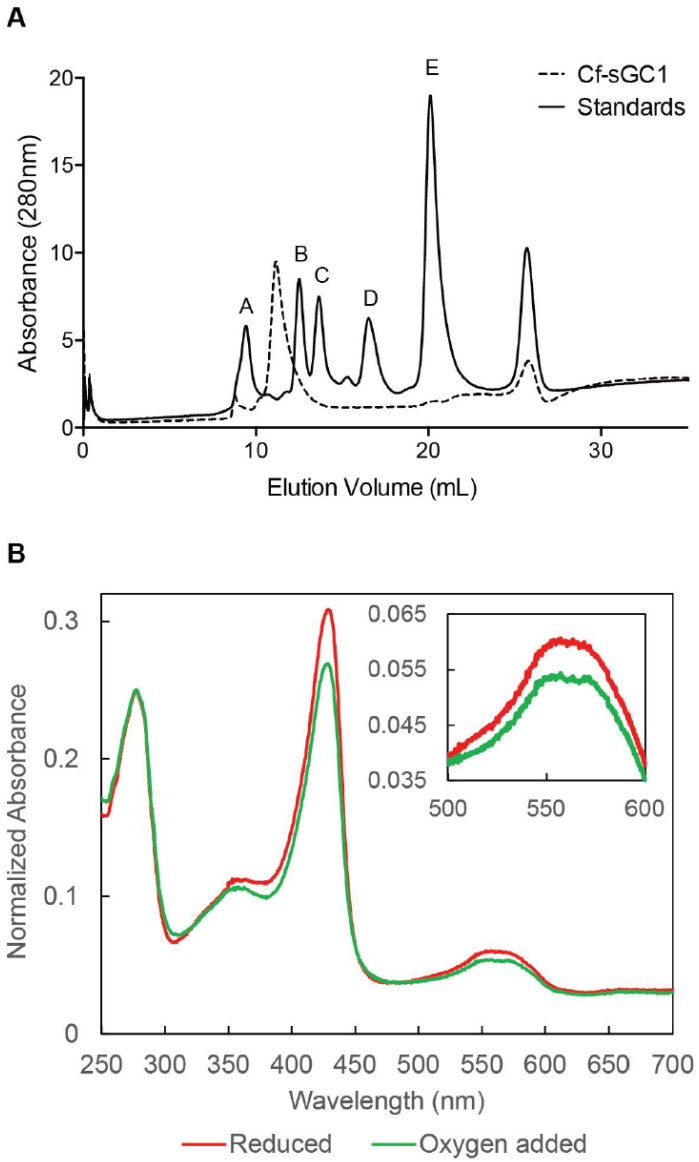


Figure A.4. Analytical size exclusion chromatography of *Cf* sGC1 and UV-vis spectrum of *Cf* sGC1 upon exposure to O_2, related to Figure 2.4. (A) Cf sGC1 eluted at ~150 kDa during the analytical size exclusion chromatography, consistent with being a homodimer of 75.7 kDa monomers. Protein standards: A, thyroglobulin, 670 kDa; B, alcohol dehydrogenase, 150 kDa; C, bovine serum albumin, 66.5 kDa; D, DNAse I, 31 kDa; E, lysozyme, 14 kDa. (B) The UV-vis spectrum of Cf sGC1 upon exposure to O2, normalized to A280. The Soret to A280 ratio decreased when oxygen was added to the sample, which could indicate oxidation of the heme and loss of the heme through the process. There was no observable formation of a low-spin 6-coordinate Fe(II)-O2 complex (as indicated by the lack of shift of the Soret peak).

Protein	Condition	Soret (nm)	α	β	References
Cf sGC1	Unliganded	429	559	559	This work
	NO	399	572	540	This work
	CO	423	568	538	This work
Rattus norvegicus sGC	Unliganded	431	562	562	(Zhao, 1997)[6]
	NO	399	572	537	(Zhao, 1997)[6]
	CO	423	568	541	(Zhao, 1997)[6]

Table A.1. *Cf* sGC1 UV-vis absorption wavelengths; *Rattus norvegicus* sGC is a well characterized, NO-selective sGC, related to Figure 2.3.

Appendix B

Supplementary methods

UV-vis spectroscopy

UV-vis absorption spectra were collected using a Cary 300 UV-vis spectrophotometer (Agilent Technologies) equipped with Cary Temperature Controller and a sample cell Peltier cooler. *Cf* sGC1 was first reduced using 5 mM sodium dithionite over 15 minutes in an anaerobic chamber (Coy) and buffer exchanged into Buffer F (50 mM HEPES, 150 mM NaCl, 5% glycerol, pH 7.4) using a pre-equilibrated Zeba spin desalting column (Thermo Scientific). Fe(II)-NO bound *Cf* sGC1 was generated through adding 100 μM proliNONOate. The NO-bound Cf sGC1 sample was buffer exchanged into buffer F again to remove excess NO. NO dissociation under aerobic conditions was monitored at 4 °C over 12 hours through UV-vis spectra collected at 20-minute intervals.

Heme reconstitution

Heme reconstitution of *Cf* sGC1 was carried out as previously reported.[7] Briefly, prior to reconstitution, *Cf* sGC1 was reduced using 5 mM DTT, and the protein concentration was quantified using UV-vis absorbance. A DMSO solution of hemin chloride was added to a final concentration of 1x and 2x molar equivalents to homodimeric *Cf* sGC1, and the protein was incubated at 25 °C for 15 minutes. Protein was then buffer exchanged to an Ar-sparged buffer F (50 mM HEPES, 150 mM NaCl, 5 % glycerol, pH 7.5, 0.22 μm filtered) using a pre-equilibrated Zeba spin desalting column (Thermo Scientific). UV-vis absorption spectra were collected on a NanoDrop 2000 spectrophotometer (Thermo Scientific), and the specific activity of the protein sample was quantified.

Intact protein mass spectrometry

Purified proteins were buffer exchanged into 25 mM ammonium bicarbonate buffer, pH 7.5 using Zeba Spin columns (Thermo Scientific). Subsequently, protein concentration was quantified by UV-vis absorbance and diluted to a final concentration of 5 μM. Prior to injection, samples were centrifuged at 4 °C, 21,130 x g for 10 minutes, and the supernatant was collected for analysis. Liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) was carried out on an Agilent 1200 Series LC system coupled to an Agilent 6224 time-of-flight mass spectrometer in positive ion mode. Chromatographic separation of protein sample was achieved over a Proswift RP-4H (1.0 mm × 50 mm, monolithic phenyl, Dionex) column using a mobile phase composed of 99.9% water, 0.1% formic acid v/v (MS solvent A) and 99.9% acetonitrile, 0.1% formic acid v/v (MS solvent B). The elution gradient was developed over 8 minutes, 5% to 100% B at 0.3 mL/min. Data were collected and deconvoluted using Agilent MassHunter software. Mass spectra were visualized using the open-source online tool Chartograph (chartograph.com).

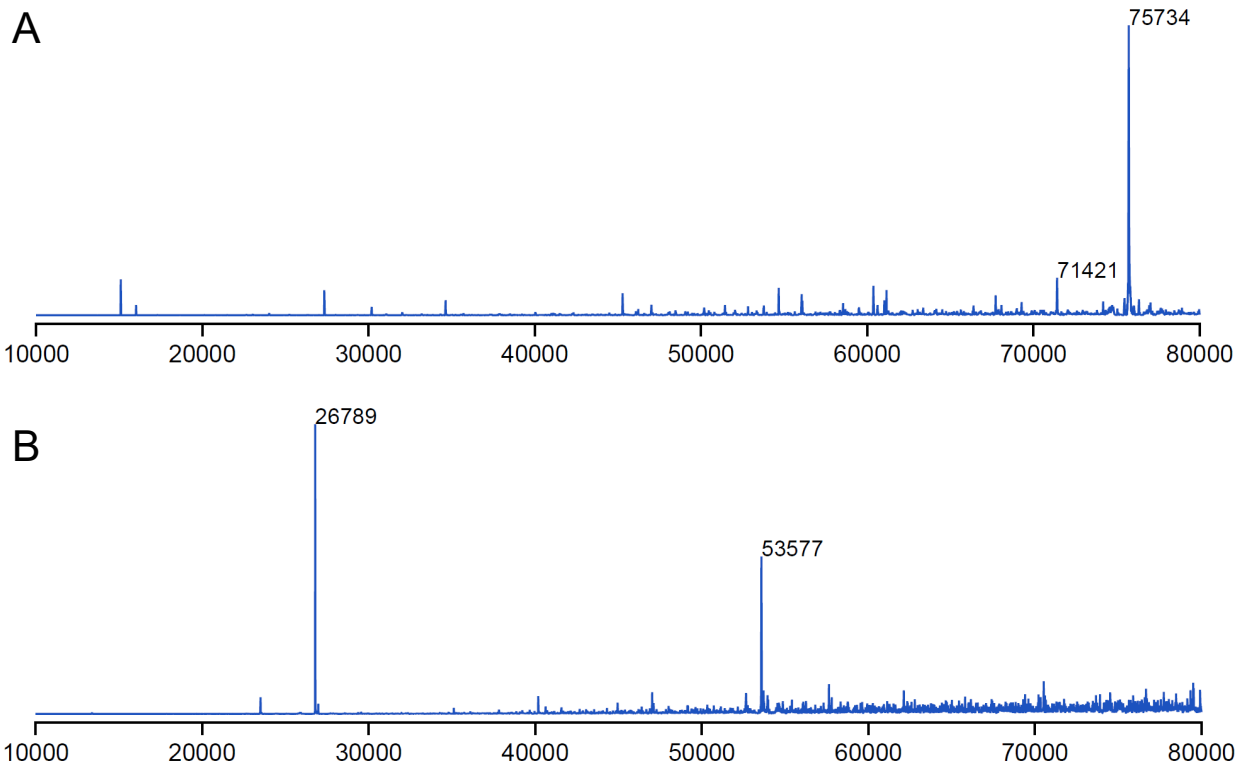


Figure B.1. Mass spectra of (A) *Cf* sGC1 and (B) sGC1-CAT. Expected molecular weight: *Cf* sGC1: 75732.3 Da, sGC1-CAT: 26789.5 Da.

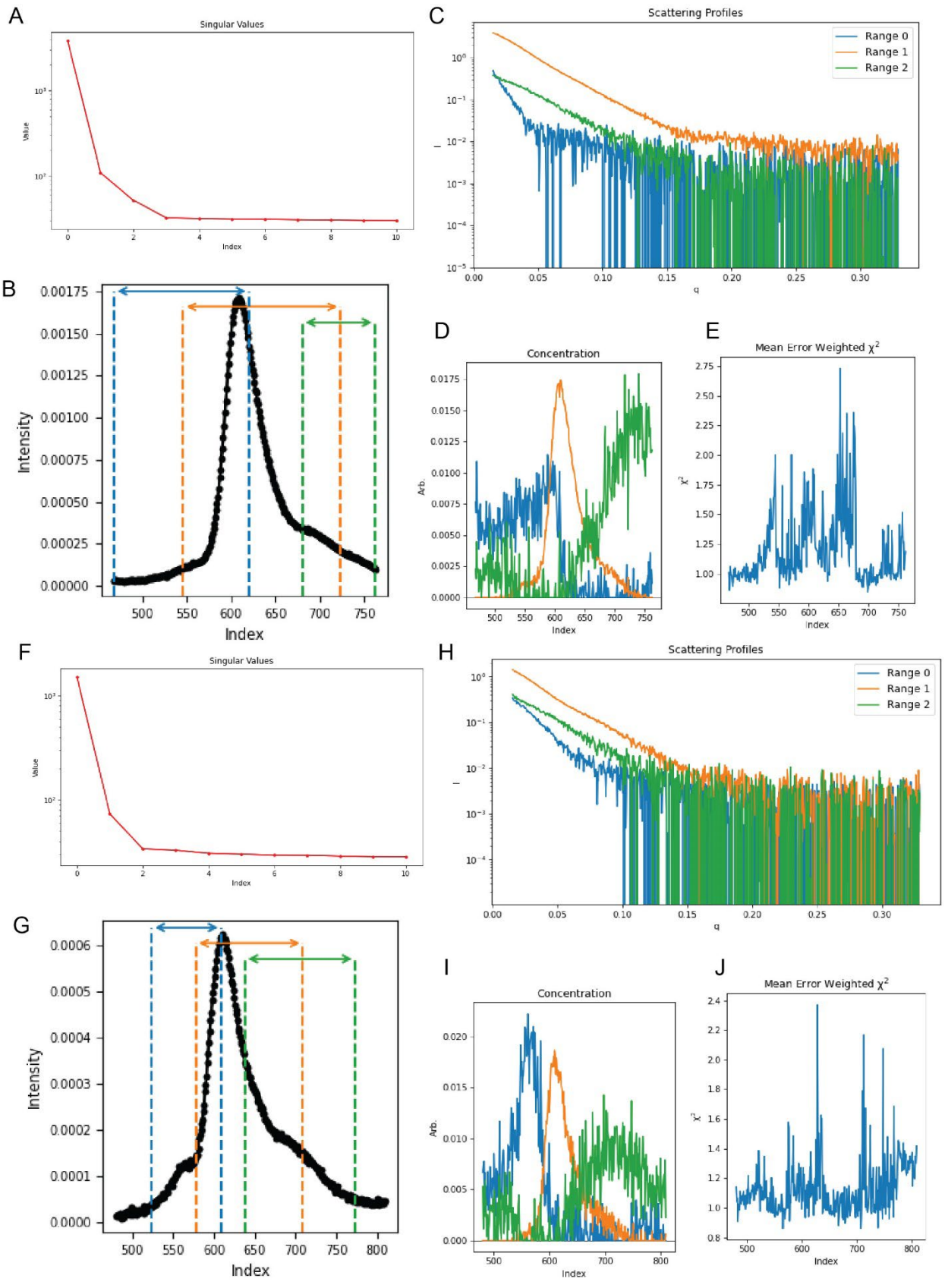

A
Singular Values
Value
Index
B
Intensity
Index
C
Scattering Profiles
Range 0
Range 1
Range 2
I
q
D
Concentration
Arb.
Index
E
Mean Error Weighted χ^2
χ^2
Index
F
Singular Values
Value
Index
G
Intensity
Index
H
Scattering Profiles
Range 0
Range 1
Range 2
I
q
I
Concentration
Arb.
Index
J
Mean Error Weighted χ^2
χ^2
Index

Figure B.2. Evolving factor analysis (EFA) calculations for *Cf* sGC1 under unliganded and NO-bound conditions**.** (A)-(E), EFA results of unliganded *Cf* sGC1. (F)-(J) EFA results of NO-bound *Cf* sGC1. (A) and (F), to carry out EFA, singular value decomposition (SVD) was used to determine that there are three significant singular values. Then, EFA was used as implemented in RAW to determine each component.[8,9] (B) and (G), elution profiles are shown to indicate the boundaries of each component. (C) and (H), SAXS scattering profiles represent the scattering of the species corresponding to each component and represent one of the singular vectors used in the EFA analysis.[10] (D) and (I) are the graphs of the second singular vector that represent concentration, and each differently colored curve represents a distinct component determined by EFA calculations. The position of the peak on the X-axis describes the retention time of the component as determined by EFA. (E) and (J), the mean error χ^2 is a measure of the goodness of the EFA calculation.

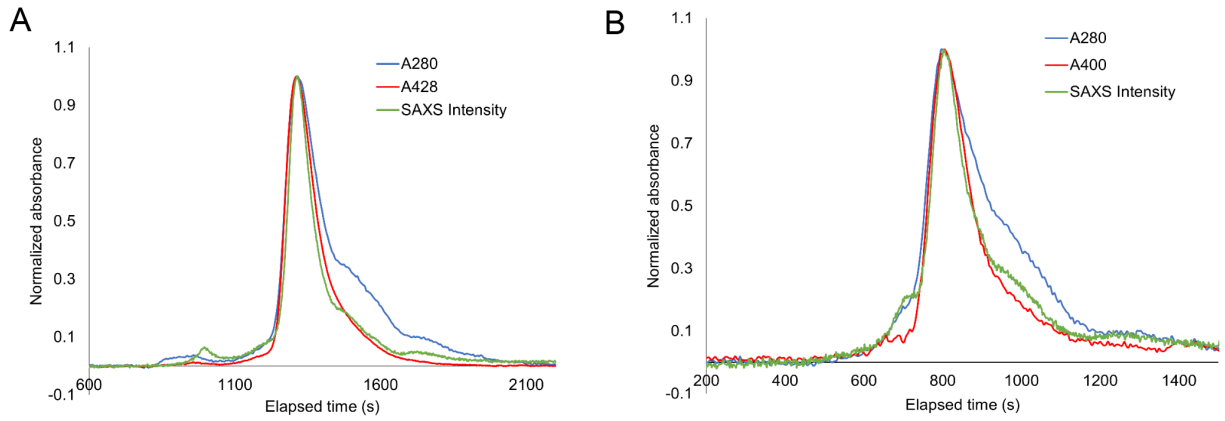


Figure B.3. Comparison of *Cf* sGC1 sample elution using UV-vis absorbance or SAXS scattering intensity. (A) Chromatogram of *Cf* sGC1 elution under Fe(II)-unliganded condition revealed a species that elutes following the major peak. This species is likely heme-free, as monitored by UV-vis absorption at 428 nm. In (B) chromatogram of *Cf* sGC1 elution under Fe(II)-NO condition a similar heme-free species was also observed, using Fe(II) heme-NO absorption at 400 nm for monitoring.

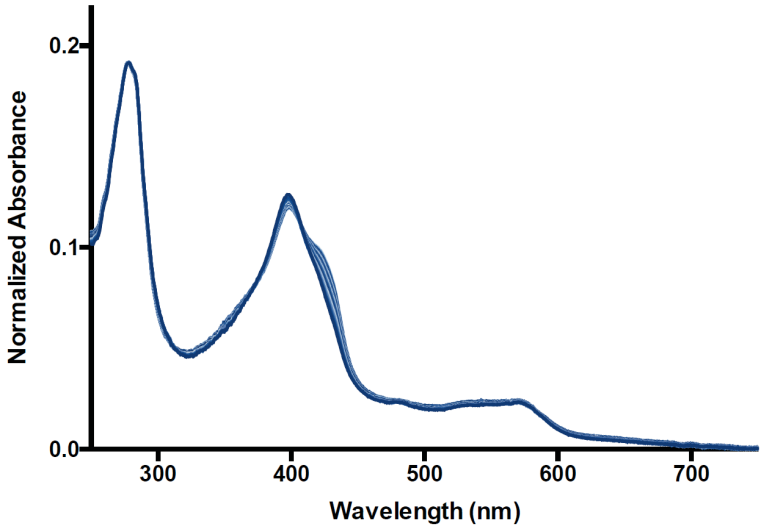


Figure B.4. Dissociation of NO from the Fe(II) – NO complex of *Cf* sGC1 over a period of 12 hours. At 4 °C, the Fe(II)-NO complex of *Cf* sGC1 was stable over the duration of the activity assay, with an estimated half-life greater than 12 hours.

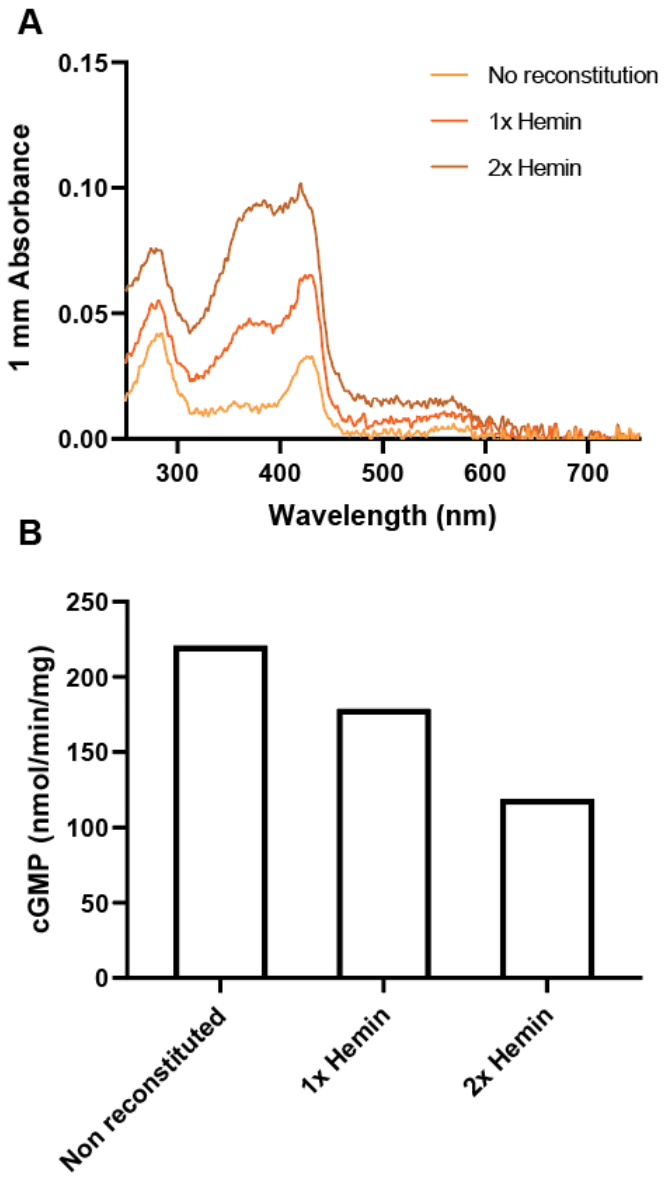


Figure B.5. Reconstitution of *Cf* sGC1. When *Cf* sGC1 was treated with an excess of hemin, the absorbance at 350 nm increased, indicative of non-specifically bound heme. When reconstituted sGC was treated with an excess of NO and its specific activity was measured, a lowered activity was observed. This suggests that the non-specifically bound heme interfered with normal function of the protein.

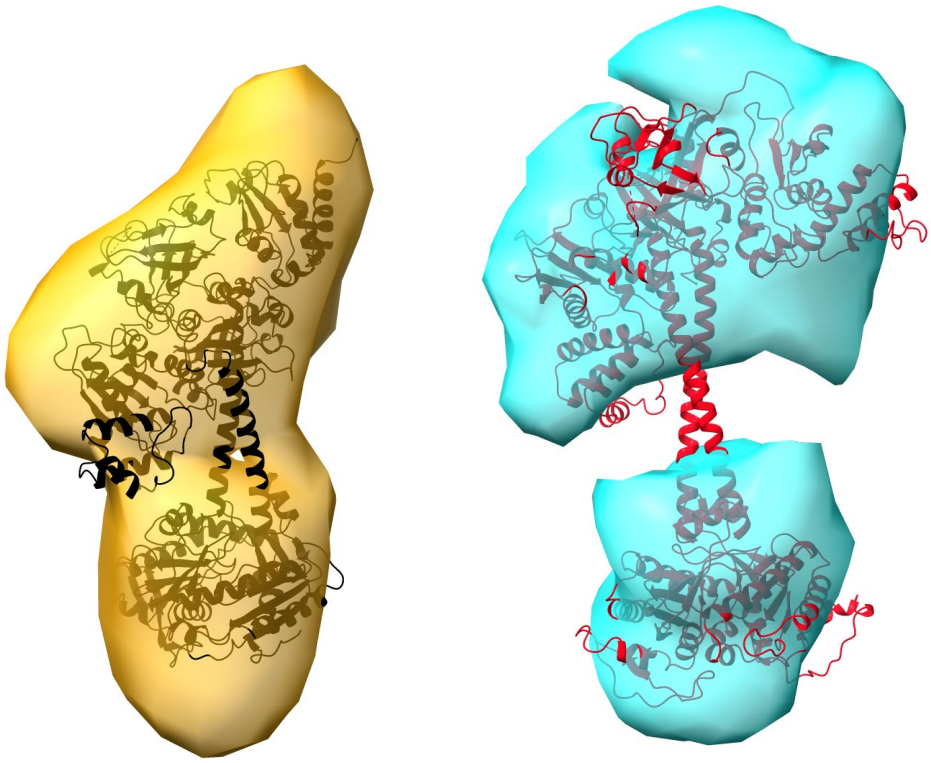


Figure B.6. DENSS[11] *ab initio* 3D reconstruction of the structure of *Cf* sGC1 reveals a more extended conformation in the NO-present condition (Cyan) when compared to the NO-free (Gold) condition. Alignment of the homology model of *Cf* sGC1 (Swiss-Model) built using human sGC inactive form (black) and active form (red) demonstrates that the NO-free conformation more resembles the inactive form of α/β-type sGC, and the NO-present conformation more resembles the active form of α/β-type sGC.

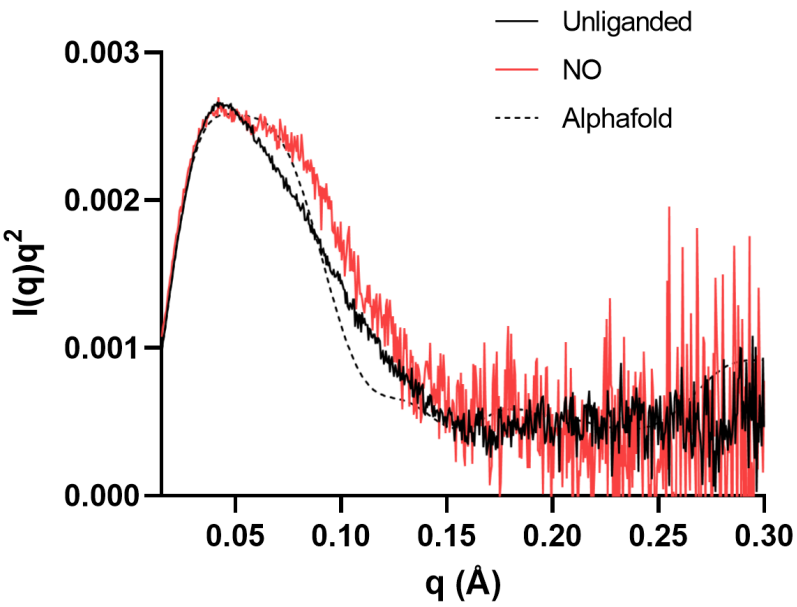


Figure B.7. Experimentally-determined SAXS profile of *Cf* sGC1 does not align well with the predicted SAXS profile of the AlphaFold-predicted structure, suggesting a conformational asymmetry akin to that seen in α/β-type sGC proteins. The predicted SAXS profile was calculated using CRYSOL.[12]

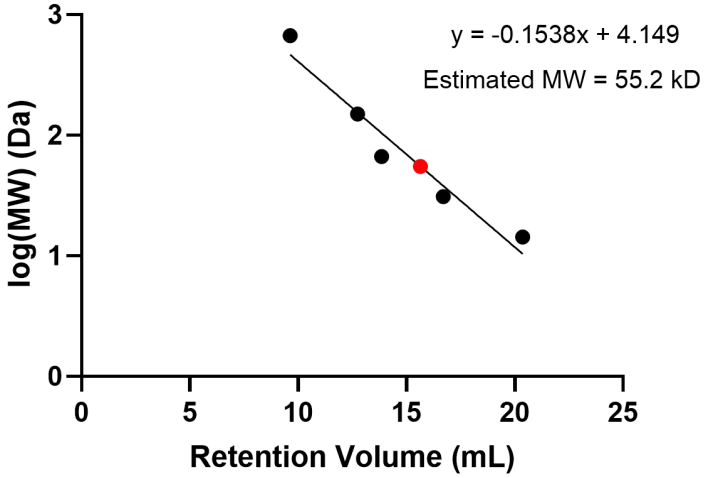


Figure B.8. Molecular weight estimation of sGC1-CAT based on the retention volume and comparison to a molecular weight standard composed of thyroglobulin (670 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66.5 kDa), DNAse I (31 kDa) and lysozyme (14 kDa). The logarithmic of the molecular weight was plotted on the y-axis and fitted to a linear curve. The linear fit was used to calculate the estimated molecular weight of sGC1-CAT of 55.2 kDa.

Table B.1. DNA sequences of proteins used in this study

Protein	Sequence
Cf sGC1	ATGTATGGCTTGGTGCACGAAGCGCTCGTGCGCGTCATCTTGAGCTTCG AAGATGGCAGCAAGCACCTACAGCAGATTGTGGAGAGAGCTGGCCTGC AGGACATGGACATCATCAATAATCACACAAACTATGATGATGAAACAAC GATGCACGTGATCGAGATCGCATGTGAAGTCATGGGAATACAAGCCGAG GAAGCTCTCACTTTGCTTGGCCCAGCTTTTGTGGAGCTTGCAATAGAGC GAGGATATCAGCCTATGTTGGAATCCCTGGGAAATTCGTTCTATGAATTT GTTGATAATTTGGACAACTTGCACCAAAACTTGGCAATGCTTTACCCAG GCATGAAGGCGCCGAGCTTTCGCTGCTCTGAAAGGCTAGACGGCACTG GGCTGGAGCTGTATTACTACAGTACTCGAGCTGGTCTTGGACCCTTTGTC ATCAGCCTGATGATCTACATTGGTGACAAGCTGTTTGAAACTGTCGTGG ACATGACAGTCCTTGAGCGTCGCGAAGACGGTGCTCGCTGCGATGTTTT CCTGATGCAGCTGCCTCAACACAAAGGAGACATTGGACCTTTATCCTCG AAGCCATCTGTTGACTACTCTCTAACCGTGTCTGCTCTCGAGCCATCGAT CGTGGCTGGCTTATTCCCGTGGCATATCGAGCTGGACGAAGACTTGATG GTGACCAGCATTGGCTCTGCGCTTCGTCGGATGCTGCCGCCGGAGATGC ACATAGGCCTGAACTTTAGGGACCTGGGTCGCGTTGTTCGCCCACTGCT ATCCCATCAAACGTTTGCTGGCATTTGCGAGCATGCCAATGCCGCCTTCT TGCTCGAAATCAAAGGAACATCCTTCAAGACACGCGCTCGTGGGGGCA ACGTACTCAGCCGCCAGTCTGTGTCGTCCACGGGGTCCGCCGCCGCTCC TTGCCCTTTTGCAAACAAAATGTTCTCTGCCGCCTCTATGGCAAGCATCA CCAGCCTGCTCGAGTCTGATGATTACCTTAAGCTCAAGGGGCAGATGGT GAAGATCAGGCACAATCGCGTTCTGTTTGTCTGCCTACCGAACGTGCGC GGGATCAATGAGATGAGCGAGCGCGGTATCAAGCTTGCCGACATCCCCG TCTTCACCACGGCTCGAGACCTTATTTTGACGGCCGACCACCAGCTGGC GACATTGAATATGGCTGCGCAGCTGCAGGAGACGACAGCCTTCCTTGAT CGAGCACTTGCTGACGTTGAAGTGGAAAAGGAGCGTGTGCAGTCGTTG CTTCACAGCATCTTGCCATCGAGTATCGCGGAGAAGCTCGCCAACGGAG AGACTGTTGAGCCATGCGAATATCCCAATTGCGCCATGCTCTTCAGCGAT ATTTGCTCCTTCACGCAGATGAGCAGTCAGGTCACACCGAAACAGGTC ATGAGCATGTTGGATGCGCTGTTCCGCGGACTAGATGACCTGTGCATCT CGCACGATGTGTACAAGTACGAGACGATTGGTGATTGCTTTGTAGTAGC CTCAGGTGTGCCGATTGAAGATCGGCGCTATGCTTCCAAGCTGGCTGCC TTTGCCATTGACTTGATCAAGCACAGTCGCAGCATCCTGTCTCCGTTGG ATGGCGAGCCTATCTCGGTTCGATGCGGCATGAACGCGGGCCCAGTGAC AACCGGTGTGATTGGTCGCGATCGCCCTCGGTACAACATCTTTGGCGAC ACTGTCAATACTGCCTCGCGCATGGAAAGCAACGGTGTTCCCAATTGCA TCCAGGTATCGCTGGCGATGAAGGAGGCAATCGAAAGTGTGGACAAGA CGTTCCGCTTCAAGGAGCGTTACGGAGGTGTTGAGGCCAAGGGCAAGG GTCGTCTGCGATCTTTCTTCATCGTCGGTCGACATGGGATCGATCCTAGC TTGCTGCCTCCCGAGGAGTATCTTGCGCCTCCGGTGGAACCCCAGCCCG AACCAGTCCCAACGCCCTCACCTAGACCTCGGCGTCGCAATCCTCTTGT CAACGTTGGCAAGCAGACTATAGTTCATCACCATCATCACCATTAA

sGC1-CAT	ATGGAGCCATGCGAATATCCCAATTGCGCCATGCTCTTCAGCGATATTTG CTCCTTCACGCAGATGAGCAGTCAGGTCACACCGAAACAGGTCATGAG CATGTTGGATGCGCTGTTCCGCGGACTAGATGACCTGTGCATCTCGCAC GATGTGTACAAGTACGAGACGATTGGTGATTGCTTTGTAGTAGCCTCAG GTGTGCCGATTGAAGATCGGCGCTATGCTTCCAAGCTGGCTGCCTTTGC CATTGACTTGATCAAGCACAGTCGCAGCATCCTGTCTCCGTTGGATGGC GAGCCTATCTCGGTTCGATGCGGCATGAACGCGGGCCCAGTGACAACC GGTGTGATTGGTCGCGATCGCCCTCGGTACAACATCTTTGGCGACACTG TCAATACTGCCTCGCGCATGGAAAGCAACGGTGTTCCCAATTGCATCCA GGTATCGCTGGCGATGAAGGAGGCAATCGAAAGTGTGGACAAGACGTT CCGCTTCAAGGAGCGTTACGGAGGTGTTGAGGCCAAGGGCAAGGGTCG TCTGCGATCTTTCTTCATCGTCGGTCGACATGGGATCGATCCTAGCTTGC TGCCTCCCGAGGAGTATCTTGCGCCTCCGGTGGAACCCCAGCCCGAACC AGTCCCAACGCCCTCACCTAGACCTCGGCGTCGCAATCCTCTTGTCAAC GTTGGCAAGCAGACTATAGTTCATCACCATCATCACCATTAA

Appendix C

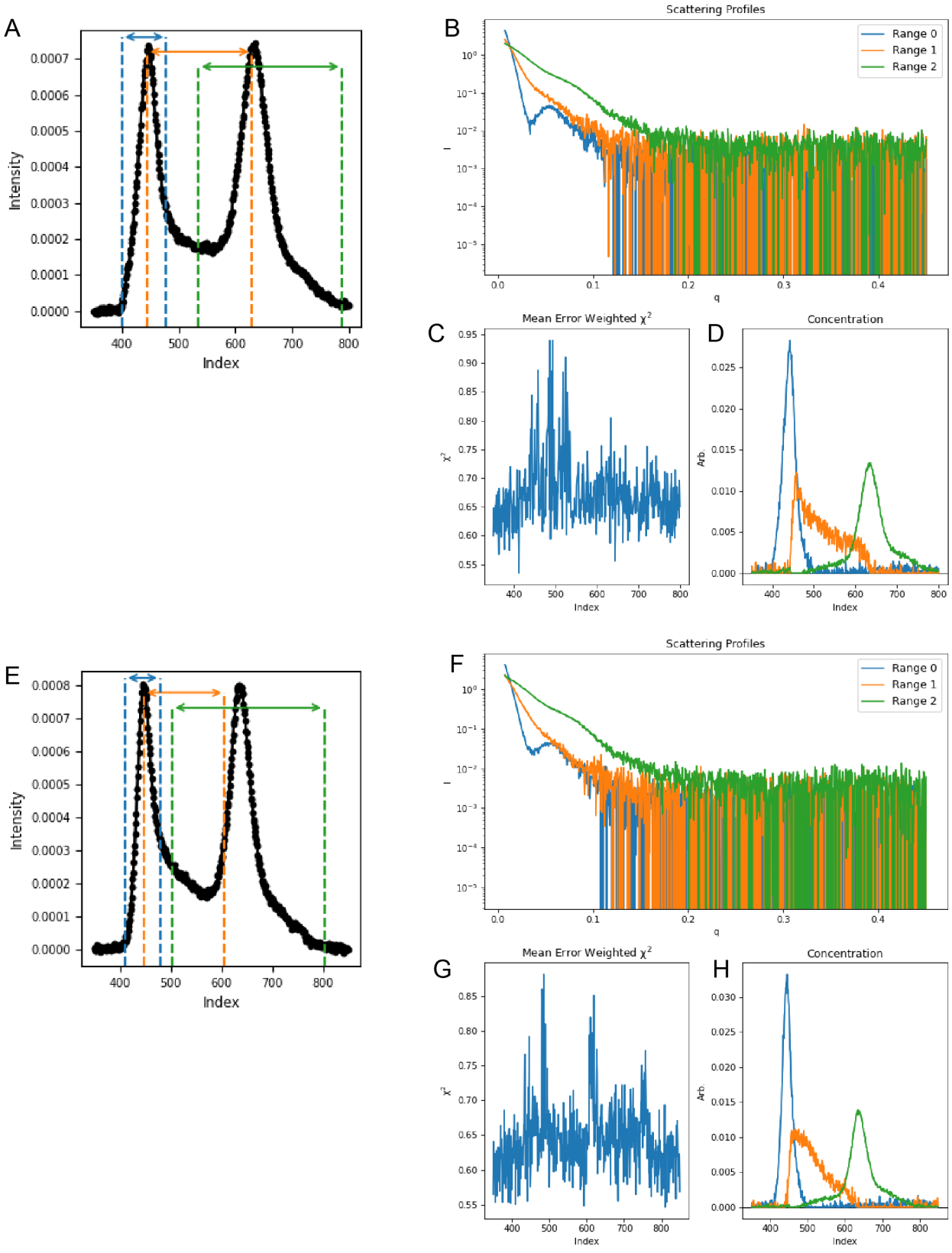

A
Intensity
Index
B
Scattering Profiles
Range 0
Range 1
Range 2
I
q
C
Mean Error Weighted χ^2
χ^2
Index
D
Concentration
Arb.
Index
E
Intensity
Index
F
Scattering Profiles
Range 0
Range 1
Range 2
I
q
G
Mean Error Weighted χ^2
χ^2
Index
H
Concentration
Arb.
Index

Figure C.1. Evolving factor analysis (EFA) calculations for *Cf* sGC4 under unliganded and NO-bound conditions. (A)-(D), EFA results of unliganded *Cf* sGC4. (E)-(H) EFA results of NO-bound *Cf* sGC4. To carry out EFA, singular value decomposition (SVD) was used to determine that there are three significant singular values. Then, EFA was used as implemented in RAW to determine each component.[10] (A) and (E), elution profiles are shown to indicate the boundaries of each component. (B) and (F), SAXS scattering profiles represent the scattering of the species corresponding to each component and represent one of the singular vectors used in the EFA analysis. (C) and (G), the mean error χ^2 is a measure of the goodness of the EFA calculation. (D) and (H) are the graphs of the second singular vector that represent concentration, and each differently colored curve represents a distinct component determined by EFA calculations. The position of the peak on the X-axis describes the retention time of the component as determined by EFA.

A

B

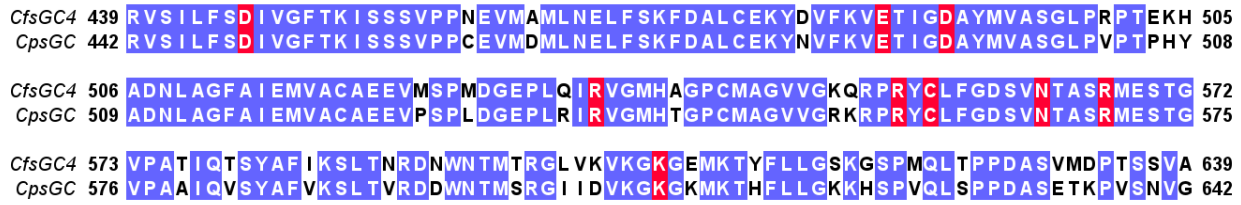


Figure C.2. Sequence alignment of *Cf* sGC4 to *Cp* sGC1 in the (A) H-NOX domain and (B) the CAT domain. Residues highlighted in red are key residues for (A) heme binding and (B) GTP binding and catalysis. At the conserved tyrosine position, *Cp* sGC1 has a phenylalanine, suggesting that *Cp* sGC1 is NO selective. *Cp* sGC1 also has all necessary catalytic domain residues consistent with its prediction function as a guanylate cyclase.

References

(1) Torruella, G.; de Mendoza, A.; Grau-Bové, X.; Antó, M.; Chaplin, M. A.; del Campo, J.; Eme, L.; Pérez-Cordón, G.; Whipps, C. M.; Nichols, K. M.; Paley, R.; Roger, A. J.; Sitjà-Bobadilla, A.; Donachie, S.; Ruiz-Trillo, I. Phylogenomics Reveals Convergent Evolution

of Lifestyles in Close Relatives of Animals and Fungi. *Curr. Biol.* **2015**, *25* (18), 2404–2410

(2) Huang, S. H.; Rio, D. C.; Marletta, M. A. Ligand Binding and Inhibition of an Oxygen-Sensitive Soluble Guanylate Cyclase, Gyc-88E, from *Drosophila*. *Biochemistry* **2007**, *46* (51), 15115–15122

(3) Boon, E. M.; Marletta, M. A. Ligand Discrimination in Soluble Guanylate Cyclase and the H-NOX Family of Heme Sensor Proteins. *Curr. Opin. Chem. Biol.* **2005**, *9* (5), 441–446

(4) Boon, E. M.; Huang, S. H.; Marletta, M. A. A Molecular Basis for NO Selectivity in Soluble Guanylate Cyclase. *Nat. Chem. Biol.* **2005**, *1* (1), 53–59

(5) Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J.-F.; Guindon, S.; Lefort, V.; Lescot, M.; Claverie, J.-M.; Gascuel, O. Phylogeny.Fr: Robust Phylogenetic Analysis for the Non-Specialist. *Nucleic Acids Res.* **2008**, *36* 465–469

(6) Zhao, Y.; Marletta, M. A. Localization of the Heme Binding Region in Soluble Guanylate Cyclase. *Biochemistry* **1997**, *36* (50), 15959–15964

(7) Horst, B. G.; Stewart, E. M.; Nazarian, A. A.; Marletta, M. A. Characterization of a Carbon Monoxide-Activated Soluble Guanylate Cyclase from *Chlamydomonas reinhardtii*. *Biochemistry* **2019**, *58* (17), 2250–2259

(8) Hopkins, J. B. *BioXTAS RAW 2* : New Developments for a Free Open-Source Program for Small-Angle Scattering Data Reduction and Analysis. *J. Appl. Crystallogr.* **2024**, *57* (1), 194–208

(9) Hopkins, J. B.; Gillilan, R. E.; Skou, S. *BioXTAS RAW* : Improvements to a Free Open-Source Program for Small-Angle X-Ray Scattering Data Reduction and Analysis. *J. Appl. Crystallogr.* **2017**, *50* (5), 1545–1553

(10) Meisburger, S. P.; Taylor, A. B.; Khan, C. A.; Zhang, S.; Fitzpatrick, P. F.; Ando, N. Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-Ray Scattering. *J. Am. Chem. Soc.* **2016**, *138* (20), 6506–6516

(11) Grant, T. D. Ab Initio Electron Density Determination Directly from Solution Scattering Data. *Nat. Methods* **2018**, *15* (3), 191–193

(12) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. FoXS, FoXSDock and MultiFoXS: Single-State and Multi-State Structural Modeling of Proteins and Their Complexes Based on SAXS Profiles. *Nucleic Acids Res.* **2016**, *44* (W1), W424–W429