SUPPORTING INFORMATION FOR:

Interfacial Residues Promote an Optimal Conformation of the Catalytic Center in Human Soluble Guanylate Cyclase: Heterodimerization Is Required But Not Sufficient For Activity

Franziska Seeger\textsuperscript{1,3}, Royston Quintyn\textsuperscript{2}, Akiko Tanimoto\textsuperscript{2}, Gareth J. Williams\textsuperscript{3}, John A. Tainer\textsuperscript{3,4}, Vicki H. Wysocki\textsuperscript{2}, and Elsa D. Garcin\textsuperscript{1,*}

SUPPLEMENTARY RESULTS

TEV Cleavage of αβGC Yields Two Populations of αGC Subunits with the Minor One Being Favored for Crystallization. The N- and C-terminal residues of αGC and βGC are not visible in the crystal structure of the αβGC heterodimer. To rule out the possibility that these proteins may be proteolysed or degraded, and to unambiguously determine the masses of the purified proteins, we performed nano-ESI/MS analysis on the αGC and βGC proteins purified individually or co-purified (summarized in Table 2). Nano-ESI/MS analysis of αGC revealed a mixture of dimeric αGC proteins composed of a major species with a measured mass of 49,489 Da ± 20.3 Da and a minor species with a measured mass of 43,500.1 Da ± 55.4 Da (Figure S1A). While surprising, this result was not completely inexplicable. Closer inspection of the amino-acid sequence revealed that αGC contains not only the engineered TEV cleavage site between residues 465 and 466 (N-terminal site), but also an endogenous TEV cleavage site between residues 662 and 663 (C-terminal site) that was not described before\textsuperscript{1}. Thus, the nano-ESI data suggested that TEV cleavage of αGC yielded a mixture of αGC(466-690) and αGC(466-662) proteins. Cleavage at the C-terminal site was relatively inefficient, as the spectrum showed about 79% of αGC(466-690) and 21% of αGC(466-662). Furthermore, we were still able to detect αGC
by western blotting with an antibody recognizing the C-terminal region 673-690, thus confirming that most of the protein was not cleaved at the C-terminal TEV site during our purification protocol.

In parallel, inspection of crystal packing for our structure showed that the 28 missing C-terminal αGC residues (663-690) could not fit in the available space between symmetry mates. Taken together, these results suggested that we had crystallized the heterodimeric αβGC protein with a truncated αGC(466-662) subunit. In contrast, a similar analysis of the crystal packing in the β₁β₂GC structure (PDB entry 2WZ1) shows that the N- and C-terminal residues could easily fit in the available space between symmetry mates. Overall, these results likely explain our high success rate in crystallizing the ββGC homodimer compared to the αβGC heterodimer.

**Comparison with the Mutant Heterodimeric αβGC Structure.** We superimposed our structure to that of the mutant heterodimeric αβGC recently solved.¹ First, the two subunits in the mutant heterodimer are closer to each other in the region encompassing βGC helices α1 and α2, and αGC N-terminal β-strand β1 (rmsd=1.3 Å). This difference can be explained by the engineered disulfide in the mutant structure, even though a disulfide bond was not observed in the mutant x-ray structure. Indeed, x-ray-induced disulfide bond cleavage in protein crystals is very common for data collected at synchrotron sources.²,³ The fact that a disulfide was formed in solution and then cleaved during data collection is supported by the observation that heterodimers were the predominant species only in the case of the mutant protein, while wild-type catalytic domains yielded mostly ββGC homodimers in previous studies.¹ Regardless, βGC Cys476 (engineered) and αGC Cys595 (endogenous) are in close proximity (3.8 Å) in the mutant structure and a covalent connectivity could indeed pull the βGC and αGC subunits closer to each other. Second, the two subunits in the mutant structure are farther from each other in the region
Interfacial residues promote an optimal alignment of the catalytic center in sGC

encompassing βGC helix α4, αGC helix α1, and the αGC substrate loop (rmsd=0.6 Å). This difference is likely caused by the αGC substrate-binding region (residues 525-532) adopting a more extended conformation that pushes the βGC subunit away in the mutant structure.

**Analysis of the Heterodimeric (αβGC) and Homodimeric (β1β2GC and ααGC) Interfaces.** We analyzed the dimer interfaces in the homodimer (PDB entry 2WZ1) and the heterodimer (our structure) catalytic domains with PISA. In the homodimeric structure, the dimer interface involves 163 atoms (49 residues), and 159 atoms (44 residues) for β2GC and β1GC, respectively. It buries a surface area of 1,455 Å$^2$ and comprises 12 hydrogen bonds, 1 salt bridge, and numerous hydrophobic interactions (Table S1). In the heterodimeric structure, the dimer interface involves 126 atoms (44 residues), and 145 atoms (41 residues) for βGC and αGC, respectively. It buries a surface area of 1,312 Å$^2$, and comprises nine hydrogen bonds, two salt bridges, and numerous hydrophobic interactions (Figures 3C-F).

To identify potential drivers of homodimerization versus heterodimerization, we further extended our analysis to the ααGC homodimer for which no crystal structure is available. We used SWISSMODEL to model homodimeric ααGC based on the αβGC heterodimer and analyzed the interface with PISA. The ααGC interface buries 1,250 Å$^2$ and contains hydrophobic interactions, seven hydrogen bonds, and no ionic interaction. Compared to the ββGC and the αβGC dimers, the ααGC interface appears to have the fewest interactions.

We tested this prediction by size-exclusion chromatography and native ESI-MS. The independently purified αGC and βGC proteins eluted differently from a gel filtration column. The βGC protein eluted as one single sharp peak corresponding to dimeric protein, while αGC eluted as a mixture of dimeric (peak) and monomeric (shoulder) protein (Figure S2A).
Nano-ESI/MS analysis of αGC revealed a mixture of dimeric αGC proteins composed of a major species with a measured mass of 49,489 Da ± 20.3 Da and a minor species with a measured mass of 43,500.1 Da ± 55.4 Da (Figure S1A). We also observed a species with a measured mass of 24,740.8 Da ± 17.9 Da, which is close to the calculated molecular weight for monomeric αGC(466-690).

Nano-ESI/MS analysis of βGC yielded a mixture of oligomeric species. The major species had a measured mass of 49,636.7 Da ± 36.6 Da, which is close to the calculated molecular weight for the βGC(407-626) dimer. The other species observed had measured masses of 24,820.5 Da ± 47.8 Da and 99,929.9 Da ± 68.7 Da, corresponding to βGC monomer and tetramer, respectively (Figure S1B).

The nano-ESI/MS spectrum for co-purified αβGC was similar to that obtained for βGC and showed that the major species had a measured mass of 49,660.2 Da ± 38.7 Da (Figure S1C). In an effort to determine whether this species was heterodimeric αβGC, we performed Surface Induced Dissociation (SID). Nano-ESI MS/MS revealed the presence of two charge state distributions with measured masses of 24,803.9 Da ± 22.2 Da and 24,839 Da ± 53.9 Da (Figure S1F). However, we could not confidently identify this species as the heterodimer, as nano-ESI/MS/MS of the homodimers showed a similar pattern, especially for βGC (Figures S1D,E).

An area under the curve (AUC) analysis of the nano-ESI/MS spectrum (Figure S1A) confirmed the results observed with gel filtration and showed that αGC is a mixture of monomers (59%) and dimers (41%). For βGC, the nano-ESI/MS data suggested the presence of mostly dimers, with a very small population of monomers and some tetramers (Figure S1B) that were not visible by size-exclusion. This is likely due to lower sensitivity in the size-exclusion chromatography. The nano-ESI/MS spectrum for αβGC revealed a mixture of monomers (23%),
Interfacial residues promote an optimal alignment of the catalytic center in sGC
dimers (70%), and tetramers (7%), as determined by AUC analysis (Figure S1C). The major
species (dimer) was confirmed by size-exclusion chromatography (Figure S2A) while the other
species were not detectable.

The Structural Flaps Play Different Roles in the Stabilization of Homodimeric and
Heterodimeric Interfaces. Our structure allowed us to propose a major role for the structural
βGC flap located at the dimer interface in modulating the conformation of the catalytic center.
Interestingly, αGC and βGC flaps are structurally different. In the heterodimeric structure, only
the βGC flap participates in heterodimerization. We can rule out the possibility that the
conformation of the αGC flap is due to crystal packing, as the mutant heterodimeric structure
(PDB entry 3UVJ) shows an identical conformation for the αGC flap, despite different
crystallization conditions, spacegroup, and crystal packing. Instead, the structure of the
heterodimer wild-type catalytic domains suggests that the αGC flap conformation is biologically
significant, and may be important for interactions with other sGC domains.

To understand the role of these flaps in dimerization, we analyzed their conformation in
all available structures of guanylate cyclase and adenylate cyclase domains. For the guanylate
cyclase family, both heterodimeric structures (3UVJ and this study) show a flipped-out αGC flap
conformation that is not part of the dimer interface. In contrast, all homodimeric GC structures
show the “double flap-wrap” conformation stabilizing the interface (Homo sapiens, 2WZ1;
Chlamydomonas reinhardtii, 3ET6; Synechocystis PCC6803 Cya2, 2W01). For adenylate
cyclase, the type II mammalian AC structures all show the “double flap-wrap” conformation
(1AB8, 1AZS, 1CJU), regardless of activation state or oligomerization state (homodimer versus
heterodimer). Structures of homodimeric adenylate cyclases from Spirulina platensis (1WC1)
and Pseudomonas aeruginosa (3R5G) also show the “double flap-wrap”. One notable exception
Interfacial residues promote an optimal alignment of the catalytic center in sGC

is the homodimeric pH-sensing AC holoenzyme from *Mycobacterium tuberculosis* in the activated state (1Y11), where both flaps are flipped out. In this AC enzyme, the N-terminal regulatory domain auto-inhibits the C-terminal catalytic domain. X-ray structures reveal that the N-terminal α-helical domains pack snuggly on the dorsal face of the AC catalytic domain – an interaction that is allowed by the flipped out position of both flaps. Overall, this analysis suggests an important role for the flaps in regulating not only the orientation of the catalytic subunits, but also interactions with other proteins or domains.

**SUPPORTING FIGURES**

**Supporting MOVIE S1. Superimposition of wild-type and mutant αβGC catalytic domains.**

Wild-type αβGC catalytic domains are colored blue, while mutant αβGC (PDB entry 3UVJ) catalytic domains are colored orange. The view from the dorsal face shows the 3° rotation of the subunits in the wild-type catalytic domains.

**Figure S1.** Native nano-ESI/MS and SID MS/MS spectra for αGC, βGC and αβGC.

**A-C.** MS spectra were obtained by spraying 10 µM protein sample in 0.1 M NH₄OAc (pH 7.4). (A) MS spectrum of αGC shows that αGC(466-690) monomers and dimers are present, with approximately 22% present as αGC(466-662) dimers. (B, C) MS spectra for βGC and αβGC show that βGC(407-626) and αβGC exist predominantly as dimers, with a relatively small proportion of monomers and β⁴GC tetramers.

**D-F.** 10 µM of protein sample in 0.1 M NH₄OAc was sprayed, and the +13 precursor ion was chosen for MS/MS analysis at an SID voltage of 100V. (D) Dissociation of αGC(466-690) dimer produces an almost equal population of unfolded and folded monomers. (E) βGC(407-626) and (F) αβGC dissociate to produce a higher proportion of unfolded monomers:folded monomers.
**Figure S2.** Size-exclusion profiles for αGC, βGC, αβGC, and αGC661βGC suggest differences in oligomerization equilibria.

A. Protein samples were run on a Superdex 200 tricorn column at the following concentrations: 100 µM for βGC and αGC, 50 µM for αβGC. The molecular weight standards at 66 kDa and 29 kDa are indicated by arrows. The βGC and αβGC proteins run mostly as dimers, while αGC runs as a mixture of dimers (peak) and monomers (shoulder). The asymmetric shape of the αβGC elution peak suggests sample heterogeneity.

B. The αGC661βGC sample was run at 50 µM on a Superdex 75 tricorn column. The molecular weight standards at 66 kDa and 29 kDa are indicated by arrows. SDS-PAGE analysis suggests that the first peak corresponds to αGC661βGC heterodimers, while the second peak (shoulder) corresponds to ααGC homodimers.

C. SEC-MALS profile of αβGC. The % MW varies from 106% to 96% suggesting heterogeneity of the purified heterodimer.

**Figure S3.** Determination of K_D for ααGC and ββGC homodimers by nano-ESI/MS.

A. Spectra for ααGC were obtained by spraying the protein sample in 80 mM NH_4_0Ac and 20 mM TEAA (pH 7.4). Concentration of ααGC homodimers was varied by performing a serial dilution from an initial stock solution of 35 µM over a range of 0.8-30 µM.

B. Spectra for ββGC were obtained by spraying the protein sample in 80 mM NH_4_0Ac and 20 mM TEAA (pH 7.4). Concentration of ββGC homodimers was varied by performing a serial dilution from an initial stock solution of 35 µM over a range of 2-35 µM.

For all experiments, the abundance of a particular species was determined by extracting its intensity from the ion mobility mobilogram that is produced with the mass spectrum and the
AUC determined using Origin. For ββGC, low ionization efficiency of βGC at low concentrations and the presence of multiple oligomerization species prevented the exact determination of the $K_D$ for homodimerization. Thus, the $K_D$ was determined by extrapolation of the curves.
Interfacial residues promote an optimal alignment of the catalytic center in sGC

Supplementary Table 1. PISA statistics and interfacial polar interactions in αβGC, ββGC, and ααGC

| αβGC (this study) | Buried area (Å²) | ΔG(kcal/mol) | Hydrogen bonds | Salt bridges |
|------------------|------------------|--------------|----------------|--------------|
|                  | 1,334.9 (33%)    | -19.3        | 9              | 2            |
| Chain | Residue | Chain | Residue | Interaction |
|---|---|---|---|---|
| B  | Arg416 | A    | Cys494 | Hydrogen bond |
| B  | Arg416 | A    | Cys497 | Hydrogen bond |
| B  | Gly534 | A    | Asn507 | Hydrogen bond |
| B  | Lys593 | A    | Glu608 | Hydrogen bond |
| B  | Lys593 | A    | Glu608 | Hydrogen bond |
| B  | Lys593 | A    | Ser609 | Hydrogen bond |
| B  | Ser435 | A    | Lys476 | Hydrogen bond |
| B  | Val532 | A    | Asn507 | Hydrogen bond |
| B  | Glu535 | A    | Thr511 | Hydrogen bond |
| B  | Lys593 | A    | Glu608 | Salt bridge  |
| B  | Lys595 | A    | Glu608 | Salt bridge  |

| ββGC (2WZ1) | Buried area (Å²) | ΔG(kcal/mol) | Hydrogen bonds | Salt bridges |
|-------------|------------------|--------------|----------------|--------------|
|             | 1,454.8 (43%)    | -17.2        | 12             | 1            |
| Chain | Residue | Chain | Residue | Interaction |
|---|---|---|---|---|
| B2 | Asn548 | B1 | Asn431 | Hydrogen bond |
| B2 | Gly534 | B1 | Asn451 | Hydrogen bond |
| B2 | Gln535 | B1 | Thr455 | Hydrogen bond |
| B2 | Asn451 | B1 | Val532 | Hydrogen bond |
| B2 | Asn451 | B1 | Gly534 | Hydrogen bond |
| B2 | Thr455 | B1 | Gln535 | Hydrogen bond |
| B2 | Asp458 | B1 | Arg536 | Hydrogen bond |
| B2 | Glu473 | B1 | Arg539 | Hydrogen bond |
| B2 | Thr474 | B1 | Arg539 | Hydrogen bond |
| B2 | Thr474 | B1 | Arg539 | Hydrogen bond |
| B2 | Asn431 | B1 | Asn545 | Hydrogen bond |
| B2 | Glu473 | B1 | Arg539 | Salt bridge  |

ααGC (modeled in this study)

| Buried area (Å²) | ΔG(kcal/mol) | Hydrogen bonds | Salt bridges |
|------------------|--------------|----------------|--------------|
|                  | 1,238.5 (50%) | -25            | 7            | 0            |
| Chain | Residue | Chain | Residue | Interaction |
|---|---|---|---|---|
| A2 | Asn507 | A1 | Glu470 | Hydrogen bond |
| A2 | Asn507 | A1 | Val471 | Hydrogen bond |
| A2 | Val586 | A1 | Asn507 | Hydrogen bond |
| A2 | Lys476 | A1 | Cys497 | Hydrogen bond |
| A2 | Lys476 | A1 | Ser495 | Hydrogen bond |
| A2 | Asn507 | A1 | Val471 | Hydrogen bond |
| A2 | Gly588 | A1 | Asn42  | Hydrogen bond |
Interfacial residues promote an optimal alignment of the catalytic center in sGC

Supplementary sequence alignment for αGC and βGC catalytic domains

a. Sequence alignment for αGC. Numbering is indicated above the alignment for human α1GC. We used the following sequences for alignments with CLUSTALW. β-strands (arrow) and α-helices (cylinder) are indicated.

| Sequence | Species          |
|----------|------------------|
| ha1GC    | Homo sapiens α1GC (Q02108) |
| GCYA3-PANTR | Pan troglodytes (H2QQB9)       |
| GCYA3_CANFA | Canis familiaris (Q4ZHS0)     |
| GCYA3-PIG  | Sus scrofa (Q4ZHR8)            |
| GCYA1_BOVIN| Bos taurus (P19687)            |
| GCYA3-AILME| Ailurupoda melanoleuca (D2HPI9) |
| GCYA3-LOXAF| Loxodonta Africana (G3TED3)    |
| GCYA3_RAT  | Rattus norvegicus (P19686)     |
| GCYA3_MOUSE| Mus musculus (Q9ERL9)          |
| GCYA3-CRIGR| chinese hamster (G3H5I1)       |
| GCYA3_MUSPF| mustela furo (G9K3P6)          |
| GCYA3-SPETR| Spermophilus tridecemlineatus (I3M9R4) |
| GCYA3-CHICK| Pleropus alecto (L5K9T6)       |
| GCYA3-MELGA| Meleagris gallopavo (G1MRU3)   |
| GCYA3-ANOCA| Anolis carolinensis (G1KTE3)   |
| GCYA3-DANPL| Danausplexippus (G6DS1)        |
| GCYA3-APIME| Apis mellifera (Q5W7P2)        |
| GCYA3-GRYBI| Gryllus bimaculatus (F5HSB9)   |
| GCYA3-CULQU | Culex quinquefasciatus (B0WSU5) |
| GCYA3-MANSE| Manduca secta (O77105)         |
| GCYA35-CAEEL| Caenorhabditis elegans (O02298) |
| GCYA2_RAT  | Rattus norvegicus (Q9WVI4)     |
| GCYA2_HUMAN| Homo sapiens (P33402)          |

β-strands (arrow) and α-helices (cylinder) are indicated.

```
ha1GC  471  .  423

| Sequence       | Alignment |
|----------------|-----------|
| GCYA3-PANTR    | VVQAKFSNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3_CANFA    | VVQAKFSNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-PIG      | VVQAKFRNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA1_BOVIN    | AVQAKRFGNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-AILME    | VVQAKFSDTVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-LOXAF    | VVQAKFNNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3_RAT      | VVQAKFNEVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3_MOUSE    | VVQAKFSVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3_MUSPF    | VVQAKFSDVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-SPETR    | VVQAKFSNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-PTEAL    | VVQAKFSSVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-CHICK    | VVQAKFNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-MELGA    | VVQAKFNNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-ANOCA    | VVQAKFHNVTMFLSBDIVGTAICSQCSPLQVTMLNALYTRFDQQCGELDVY |
| GCYA3-DANPL    | VIEAKSHDVTMFLSBDIVGTSICATAPMVIALEDLYSVFDIFCIRLVDY |
| GCYA3-APIME    | TIEAKTPEVTMFLSBDIVGTEICTATPMVMINLQNYEQFDSCQGLDVY |
| GCYA3-GRYBI    | SIDAQTHDVTMFLSBDIVGTSICATAPMVIMLQDYSQFDMLCGQLDVY |
| GCYA3-CULQU    | QIDAKTYPDVTMFLSBDIVGTSICRAFPMVINLSELYKHDLFGFDVY |
| GCYA3-MANSE    | KIEAKSHDVTMFLSBDIVGTSICATAPMVIALEDLYSVFDIFCIRLVDY |
| GCYA35-CAEEL   | AMDAKEFADCTLFDIVFTTNICAMCTPYDVVTNLNDLYLRFDLRLVLGHDAY |
| GCYA2_RAT      | QVQARKFDVTMFLSBDIVGTAICAQCTPMQVISMNLNELYTRFDHQCGFLDY |
| GCYA2_HUMAN    | QVQARKFDVTMFLSBDIVGTAICAQCTPMQVISMNLNELYTRFDHQCGFLDY |
```
Interfacial residues promote an optimal alignment of the catalytic center in sGC
b. Sequence alignment for βGC. Numbering is indicated above the alignment for human β1GC. We used the following sequences for alignments with CLUSTALW. β-strands (arrow) and α-helices (cylinder) are indicated.

hb1GC Homo sapiens β1GC (Q02153)
GYB1-PANTR Pan troglodytes (H2QQC0)
GYB1_CANFA Canis familiaris (F1P601)
GYB1-PIG Sus scrofa (Q4ZHR7)
GYB1_BOVIN Bos taurus (P16068)
GYB1_ALME Ailurupoda melanoleuca (G1MAB2)
GYB1_LOXAF Loxodonta africana (G3TED8)
GYB1_RAT Rattus norvegicus (P20595)
GYB1_MOUSE Mus musculus (O54865)
GYB1_MUSPF mustela furo (G9K3P7)
GYB1_SPETR Spermophilus tridecemlineatus (I3M9T4)
GYB1_PTEAL Pleropus alecto (L5KBY9)
GYB1_CHICK Gallus gallus (F1P0M1)
GYB1_MELGA Meleagris gallopavo (G1MRX6)
GYB1_ANOFA Anopheles gambiae (Q7PS01)
GYB1_DANPL Danaus plexippus (G6CVW0)
GYB1_APIME Apis mellifera (Q5FAN0)
GYB1_GRYBI Gryllus bimaculatus (F5HSC0)
GYB1_CULQO Cules quinquefasciatus (B0XCR7)
GYB1_MANSE Manduca sexta (O77106)
GYB1_CAEEL Caenorhabditis elegans (Q6DNF4)
GYB2_RAT rattus norvegicus (P22717)
GYB2_HUMAN Homo sapiens (O75343)

Interfacial residues promote an optimal alignment of the catalytic center in sGC
Interfacial residues promote an optimal alignment of the catalytic center in sGC
Interfacial residues promote an optimal alignment of the catalytic center in sGC
Interfacial residues promote an optimal alignment of the catalytic center in sGC.
Interfacial residues promote an optimal alignment of the catalytic center in sGC

Figure S1

Figure S2
Interfacial residues promote an optimal alignment of the catalytic center in sGC

Figure S3

[Graph A: Relative Abundance vs. Protein Concentration for αGC monomer and αGC dimer]

[Graph B: Relative Abundance vs. Protein Concentration for βGC monomer, βGC dimer, and βGC tetramer]