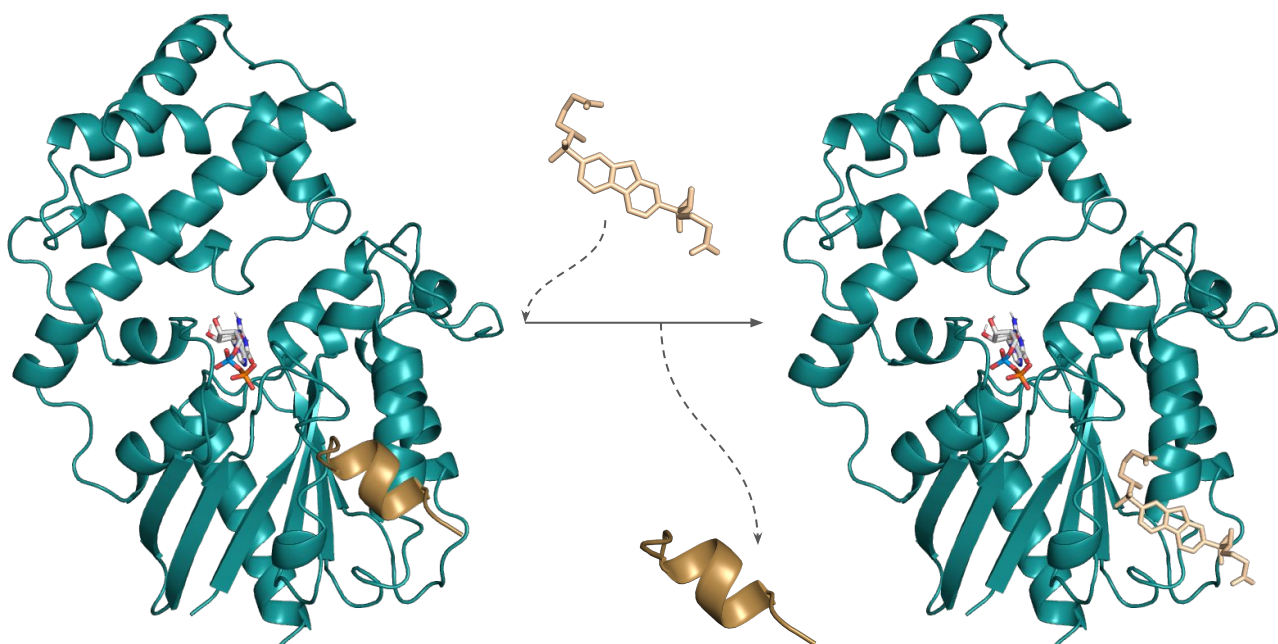


# Structure-function of the alpha subunit of the human trimeric $G_{i3}$ -protein ( $G\alpha_{i3}$ )



eman ta zabal zazu



Universidad  
del País Vasco

Euskal Herriko  
Unibertsitatea

Mariola Ferreras Gutiérrez

Ph.D. Thesis 2023



---

*I am only one, but still I am one.  
I cannot do everything, but still I can do something;  
and because I cannot do everything,  
I will not refuse to do something that I can do*

*- Helen Keller*

---

---

---

---

*We are standing on the shoulders of giants*

Cuando se usa esta metáfora, nos referimos a todos los científicos que han creado el camino que las nuevas generaciones recorreremos. Hoy, a parte de reconocer a todas las grandes mentes que hicieron posible este trabajo, quiero agradecer a mis propios gigantes personales, sin quienes nunca hubiera alcanzado esta meta.

A Paco, gracias por darme la oportunidad de trabajar bajo tu dirección, por guiarme hacia el fascinante mundo de la biología estructural, por todo lo que me has enseñado y por tu paciencia infinita. Han sido cuatro años en los que he aprendido muchísimo gracias a ti y sé que siempre los recordaré con muchísimo cariño.

A Mikel, por darme la posibilidad de formar parte de un laboratorio internacional. Gracias por dedicar tu tiempo y esfuerzo para que mi estancia con tu equipo fuese una gran experiencia.

A quienes estuvisteis conmigo en CICbioGUNE, especialmente a Nekane, Amaia y Mikel. Gracias por acompañarme en mis primeros pasos en el mundo de la ciencia y por acogerme en Bilbao.

A mis CIBeros, sois la razón por la que la ciencia sigue viva. Desde la planta 0 a la 3, sin olvidarme de Raquel, César y Sergio. Gracias a todos y cada uno de vosotros por enriquecer cada día de trabajo, por nuestros maravillosos desayunos y debates. Aunque nunca queráis venir a mi casa, siempre seréis bienvenidos (prometo no cocinar yo). Y, entrando en el 346, Antonio y Miriam, sé que será difícil (o imposible) tener mejores compañeros que vosotros y Belén, eres increíble, un claro ejemplo de lo que significa ser un científico.

A mis compañeros del GERMN junior, gracias por presentarme a una comunidad tan interesante y enriquecedora (aunque no haya ido a Jaca, tenéis que aceptarme igual).

A mi cole, por ir forjando mi camino y ayudarme a descubrir quién quiero ser y cómo lograrlo. Y, en particular, a Elena, gracias por ser mi inspiración para adentrarme en el mundo de la ciencia.

To George and Karen (and Ruby!), thank you for staying connected with me and keeping up to date with my life, even if it has been through emails, phone calls, and the occasional wonderful visits that make me feel completely at home. I owe you a thousand brunches.

---

---

A mis bioquímicos, que siempre resuelven mis dudas científicas y alimentan mi vocación. En especial a Felipe, por ayudarme a empezar, a Cris por tu eterna alegría, a David, porque hablar contigo es terapéutico, y a Mari, que podrías empezar a cobrarme de la de veces que he necesitado tu ayuda.

A los infomates, a los que tantas veces he invadido la biblioteca y cafetería. Gracias por estar presentes, por preocuparos y por los planes frikis que nunca sobran (de hecho, ¡nos faltan!).

A los Abad y Espuelas, en especial a Ana e Ignacio. Gracias por acogerme en la familia, dejarme disfrutar de planes increíbles y deleitarme con platos dignos de Masterchef cada semana.

A la chupipandi de mi padre, gracias por estar siempre pendientes desde el primer minuto. Quiero hacer mención especial a mi madrina, que vino a revolucionar Boston con su increíble cocina y sus telenovelas.

To all my Wageningen buddies - Alicia, Joram, Juanca, Kristien and Stephen - thank you for making Wageningen feel like a second home to me. Thank you for sharing with me two amazing years and for your support whenever I needed it. I truly miss living in the same city as all of you... But I love to have an excuse to travel to The Netherlands. See you in Orion!

A mis compañeras de brunch y los respectivos del CSK, por aparecer en estos últimos años y hacer que os quiera siempre a mi lado. Gracias por preocuparos, por lo atentos y detallistas que sois y por brindar una amistad increíble en la que las mimosas y Álvaro Moreno siempre serán bienvenidos.

A vosotros, los de siempre, los que estáis ahí llueva o truene, para compartir las alegrías de los mejores días y el apoyo en los peores. Gracias por esas incontables cervezas sanadoras en el bar que toque, por los viajes, por las películas malas, por las clases de pádel, por los paseos, por las falsas Nocheviejas, por los cumpleaños y las bodas. Pero también por preguntarme por mis experimentos y aguantarme cuando no salen como espero, por soportarme cuando me saturó y levantarme cuando me caigo. En definitiva, gracias por ser cada día de mi vida, la razón para vivirla. No tengo nada más que decir salvo *LGsús te da las gracias*.

---

---

A la familia más perfectamente imperfecta del planeta. A mis abuelos, sois los mayores referentes de mi vida, mi ejemplo y mis raíces. A mis tíos, por ser la alegría de cada fin de semana y por interesaros por lo que hago, aunque a veces sea difícil entenderlo. A mis primos, por ser cómplices en todos los Consejos de Nietos. A Sonia, por llegar a nuestra vida con tu alegría y vitalidad. A Aju, por ser, de lejos, la persona más leal y altruista que conozco. Gracias de verdad por ser el pilar que me sustenta, siempre me consideraré la persona más afortunada por teneros a vosotros a mi lado.

A Abad, por ser el mayor apoyo que se puede tener. Gracias por tu cariño, tu paciencia y tu increíble mente que me enamora cada día. Todo es mil veces mejor cuando estás a mi lado. No hay duda de que somos el mejor equipo del mundo.

A Pato, por ser la calma en la tormenta. Poca gente tiene la suerte de conocer a su alma gemela tan pronto, yo estuve sólo dos años sin ella. Gracias porque mi vida es mejor porque estás en ella.

Y por supuesto, gracias a las dos personas a las que les debo absolutamente todo, que siempre me han apoyado y caminado a mi lado. A mi madre, siempre fuiste, eres y serás mi persona favorita de este mundo y del siguiente. A mi padre, eres mi héroe, ojalá pueda devolverte algún día todo lo que me has dado.

Gracias a todos por esta aventura, ¡ahora toca vivir una nueva!

¡Casi se me olvida! No puedo terminar estos agradecimientos sin darles millones de gracias a Nugget y Pippin, mis pequeños gatetes que hacen que cualquier día sea diez veces mejor.

---





---

*To the giants whose shoulders I stand on*

---



---

<b>Abstract</b> .....	1
<b>Abbreviations</b> .....	7
<b>1. Introduction</b> .....	9
1.1 Guanine nucleotide-binding proteins.....	9
1.1.1 Small G proteins.....	16
1.1.2 Large heterotrimeric G proteins.....	17
1.1.2.1 G <sub>s</sub> proteins.....	21
1.1.2.2 G <sub>i</sub> proteins.....	21
1.1.2.2.1 G $\alpha$ i3.....	22
1.1.2.3 G <sub>q/11</sub> proteins.....	23
1.1.2.4 G <sub>12/13</sub> proteins.....	23
1.2 Structure of heterotrimeric G proteins.....	24
1.2.1 G $\beta\gamma$ structure.....	25
1.3 Regulation of heterotrimeric G proteins.....	26
1.3.1 Guanine nucleotide exchange factors.....	26
1.3.1.1 G protein Coupled Receptors.....	26
1.3.1.2 Non-receptor G protein GEFs.....	28
1.3.1.3 GIV is a non-receptor GEF of G $\alpha$ i3.....	28
1.3.1.4 Inhibitors of G $\alpha$ i3-GIV interaction and antimetastatic therapy.....	30
1.3.2 GDIs and GAPs.....	31
1.3.2.1 Regulator of G protein Signaling 12.....	32
1.3.2.2 Non-natural regulators.....	33
<b>2. Hypothesis and Research Objectives</b> .....	35
<b>3. Materials and Methods</b> .....	39
3.1 Protein expression and purification.....	41

---

3.1.1 Gene design for human Gai3 protein production in bacterial cells.....	41
3.1.2 Protein production and purification.....	42
3.1.2.1 Gai3 production.....	42
3.1.2.2 sGai3 production.....	43
3.1.2.3 Protein purification.....	43
3.2 Ligand sample preparation for binding studies.....	45
3.2.1 Protein fragments and non-natural sequences.....	46
3.2.2 Chemical compound.....	47
3.2.3 Gβ1γ2 dimer.....	47
3.3 Size exclusion chromatography-multiangle light scattering.....	47
3.4 Isothermal titration calorimetry.....	48
3.5 Nuclear magnetic resonance spectroscopy.....	50
3.5.1 NMR relaxation and protein backbone dynamics.....	53
3.5.2 Protein-ligand binding studies.....	55
3.5.2 Measurement of guanine nucleotide exchange rates in Gai3.....	58
3.6 Crystallography.....	59
3.7 <i>In silico</i> docking of IGGi-11 on sGai3.....	62
3.7.1 Docking with HADDOCK.....	63
3.7.2 Docking with Autodock.....	63
3.8 Cell signaling stimulation, cell lysis, and immunoblotting.....	65
<b>4. Results.....</b>	<b>67</b>
4.1 Characterize the binding of IGGi-11 to Gai3.....	69
4.1.1 GIV fragments containing the GBA motif bind similarly to Gai3.....	69
4.1.2 IGGi-11 binds to the same region of GDP-Gai3 that interacts with GIV.....	69
4.1.3 IGGi-11 specifically targets Gai3 in its GDP-bound state.....	76

---

---

4.1.4	IGGi-11's impact on Gai3-GIV1671-1705 interaction.....	77
4.1.5	IGGi-11 hinders signaling mediated by the GIV-Gai3 interaction.....	79
4.2	Crystal structure of Gai3.....	82
4.2.1	Experimental mapping of flexible regions in GDP-Gai3.....	83
4.2.2	Production and purification of sGai3.....	87
4.2.3	Suitability of sGai3 construct for crystallization assays.....	89
4.2.4	Crystallization, data collection and refinement.....	91
4.2.5	<i>In silico</i> docking of IGGi-11 with sGai3.....	96
4.2.5.1	IGGi-11 docking studies with HADDOCK.....	96
4.2.5.2	IGGi-11 docking studies with Autodock.....	98
4.3	Study the binding of other regulators to Gai3 by NMR.....	104
4.3.1	Insights into binding sites of regulators of Gai3.....	104
4.3.2	Effect of IGGI-1 on the Gai3 complexes.....	110
4.4	NMR-based measurement Gai3's nucleotide exchange rate.....	114
<b>5.</b>	<b>General Discussion</b> .....	<b>119</b>
<b>6.</b>	<b>Conclusions</b> .....	<b>125</b>
<b>7.</b>	<b>Spanish Summary</b> .....	<b>129</b>
<b>8.</b>	<b>Bibliography</b> .....	<b>137</b>
<b>9.</b>	<b>Publications</b> .....	<b>151</b>

---



# Abstract







Most human G proteins participate in cell signaling as they switch from a GDP-bound inactive state to a GTP-bound active state, and back when GTP is hydrolyzed. Dysregulation of this switching is linked to different diseases. Heterotrimeric G proteins consist of  $\alpha$ ,  $\beta$ , and  $\gamma$  subunits, and are activated by G Protein Coupled Receptors, which accelerate the guanine nucleotide exchange at  $G\alpha$ . In this way they participate in extracellular signal transduction across the cytoplasmic membrane. But some of them are also involved in intracellular signaling, activated by non-receptor proteins.  $G\alpha$ -interacting vesicle-associated protein (GIV), activates  $G\alpha i3$  and regulates cell migration. As GIV is present at high levels in metastatic cancers, disruption of the  $G\alpha i3$ -GIV interaction has potential therapeutic interest. The interaction has been previously characterized by NMR and found to be druggable.

In this thesis, the binding to GDP- $G\alpha i3$  of the small molecule IGGi-11, inhibitor of the interaction and with favorable properties to be developed into a drug, has been structurally characterized, primarily by solution NMR. The inhibitor is a competitive one, binding to the same site as a fragment of GIV does, with 10-fold lower affinity, and specific for the GDP-bound state.

High resolution structural information on the interaction would facilitate the design of IGGi-11 derivatives with higher affinity. This information is very difficult to obtain by NMR, and previous crystallization attempts were unsuccessful. The most flexible regions at the  $G\alpha i3$  chain termini were eliminated to facilitate crystallization. The trimmed protein, with the same affinity for GIV and IGGi-11 as the full-length one, crystallized, and the structure at 3.3 Å is very similar to  $G\alpha i3$  in complex with different regulators, but co-crystallization with IGGi-11 or GIV did not occur. Docking IGGi-11 indicates a heterogeneous binding within the cleft where GIV binds, and suggests that modifications breaking IGGi-11's symmetry might increase its affinity.

Binding of other regulators ( $G\beta 1\gamma 2$ , the GoLoco motif of RGS12, and non-natural peptides) was investigated by NMR. Their binding is consistent with crystal structures of homologous proteins, except that the N-terminal region of  $G\alpha i3$  interacts only weakly with  $G\beta 1\gamma 2$ . Unexpectedly, IGGi-11 at high relative concentration could bind to the G domain of the GDP- $G\alpha i3$  complexes, except the RGS12 one.

The NMR signals of the three tryptophan residues of  $G\alpha i3$  are very useful to quickly monitor binding events. Using them, the intrinsic guanosine exchange rate of  $G\alpha i3$  has been directly measured without the use of radioactivity.

---



La mayoría de las proteínas G humanas desempeñan un papel clave en la señalización celular al cambiar entre estados activo (unidas a GTP) e inactivo (unidas a GDP), regulando así procesos celulares. Esta transición desregulada está asociada con enfermedades. Las proteínas G heterotriméricas constan de subunidades  $\alpha$ ,  $\beta$  y  $\gamma$ , y se activan a través de receptores acoplados a proteínas G, participando en la transducción de señales extracelulares. Sin embargo, algunas de estas proteínas también intervienen en la señalización intracelular, siendo activadas por proteínas citosólicas. Por ejemplo, GIV activa Gai3 y regula la migración celular. Dado que GIV se encuentra en niveles elevados en cánceres metastásicos, interrumpir la interacción entre Gai3 y GIV tiene un potencial interés terapéutico y puede ser objetivo de fármacos, ya que se ha caracterizado mediante RMN.

En esta tesis, se ha caracterizado estructuralmente la unión a GDP-Gai3 a la pequeña molécula IGGi-11, un inhibidor de la interacción con propiedades favorables para su desarrollo como fármaco, principalmente mediante RMN en solución. El inhibidor es competitivo, uniéndose al mismo sitio que un fragmento de GIV, pero con una afinidad 10 veces menor, y es específico para el estado unido a GDP.

La obtención de información estructural detallada facilitaría diseñar versiones mejoradas de IGGi-11 con mayor afinidad. Sin embargo, esta información es difícil de obtener mediante RMN, y los intentos previos de cristalización no tuvieron éxito. Se eliminaron las regiones más flexibles en los extremos de la cadena Gai3 para facilitar la cristalización. A pesar de que la proteína truncada, que tiene la misma afinidad que la proteína completa por GIV e IGGi-11, cristalizó con una estructura a una resolución de 3.3 Å, no se logró la co-cristalización con IGGi-11 o GIV. Los resultados del acoplamiento computacional de IGGi-11 indican que se une de manera heterogénea en la misma hendidura donde se une GIV, y sugieren que realizar modificaciones para romper la simetría de IGGi-11 podría aumentar su afinidad.

Se investigó la unión de otros reguladores (G $\beta$ 1 $\gamma$ 2, el motivo GoLoco de RGS12 y péptidos no naturales) mediante RMN. Su unión es consistente con las estructuras cristalinas de proteínas homólogas, excepto que la región N-terminal de Gai3 interactúa débilmente con G $\beta$ 1 $\gamma$ 2. De manera inesperada, IGGi-11 en concentraciones relativas elevadas pudo unirse a los complejos GDP-Gai3, excepto al de RGS12.

Las señales de RMN de los tres residuos de triptófano de Gai3 son muy útiles para monitorear rápidamente los eventos de unión. Utilizándolas, se midió directamente la velocidad intrínseca de intercambio de guanosina de Gai3 sin necesidad de radioactividad.





# Abbreviations



---

<b>AC:</b> Adenylate cyclase	<b>IGGi:</b> Inhibitors of the GIV-G $\alpha$ i interaction
<b>AIR:</b> Ambiguous Interaction Restraint	<b>IMAC:</b> Immobilized metal affinity chromatography
<b>AlF<sub>4</sub>:</b> Tetrafluoroaluminate	<b>IPTG:</b> Isopropyl $\beta$ -D-1-thiogalactopyranoside
<b>ATP:</b> Adenosine triphosphate	<b>ITC:</b> Isothermal titration calorimetry
<b>cAMP:</b> Cyclic adenosine monophosphate	<b>K<sub>p</sub>:</b> Dissociation constant
<b>CSP:</b> Chemical shift perturbation	<b>LB:</b> Lysogeny broth
<b>DMEM:</b> Dulbecco's modified eagle medium	<b>LOPAC:</b> Library of Pharmacologically Active Compounds
<b>DSS:</b> 2,2-Dimethyl-2-silapentane-5-sulfonic acid	<b>MALDI-TOF:</b> Matrix-assisted laser desorption/ionization time-of-flight
<b>DSSP:</b> Definition of Secondary Structure of Proteins	<b>MALS:</b> Multiangle light scattering
<b>EGF:</b> Epidermal growth factor	<b>MD:</b> Molecular Dynamics
<b>EGFR:</b> Epidermal growth factor receptor	<b>NMR:</b> Nuclear magnetic resonance
<b>G protein:</b> Guanine nucleotide-binding protein	<b>NOE:</b> Nuclear Overhauser effect
<b>GAP:</b> GTPase-Activating Proteins	<b>OD<sub>600</sub>:</b> Optical density at 600 nm
<b>GBA:</b> G $\alpha$ binding and activating	<b>PDB:</b> Protein data bank
<b>GBD:</b> G-binding domain	<b>PDZ:</b> PSD-95/Discs Large/ZO-1
<b>GDI:</b> Guanine nucleotide dissociation inhibitor	<b>PI3K:</b> Phosphoinositide 3-kinase
<b>GDP:</b> Guanosine-5'-diphosphate	<b>PVDF:</b> Polyvinylidene fluoride
<b>GEF:</b> Guanine nucleotide exchange factor	<b>Ras:</b> Rat sarcoma
<b>GEM:</b> Guanine Exchange Modulator	<b>RGS:</b> Regulators of G Protein signaling
<b>GIV:</b> G $\alpha$ -interacting vesicle-associated protein	<b>RMSD:</b> Root Mean Square Deviations
<b>GPCR:</b> G protein coupled receptors	<b>RMSF:</b> Root Mean Square Fluctuations
<b>GppNHp:</b> Guanosine 5'-[( $\beta$ , $\gamma$ )-imido]triphosphate	<b>SEC:</b> Size exclusion chromatography
<b>GTP:</b> Guanosine-5'-triphosphate	<b>sGai3:</b> Short human G $\alpha$ i3
<b>GTPase:</b> GTP-binding protein	<b>SMILES:</b> Simplified Molecular Input Line Entry System
<b>GTP<math>\gamma</math>S:</b> Guanosine 5'-O-[gamma-thio]triphosphate	<b>TCEP:</b> Tris(2-carboxyethyl)phosphine
<b>HADDOCK:</b> High Ambiguity Driven Biomolecular Docking	<b>TM:</b> Transmembrane
<b>HSQC:</b> Heteronuclear single quantum coherence	<b>TROSY:</b> Transverse relaxation optimized spectroscopy
	<b>YEP:</b> Yeast extract peptone

---





# Introduction





---

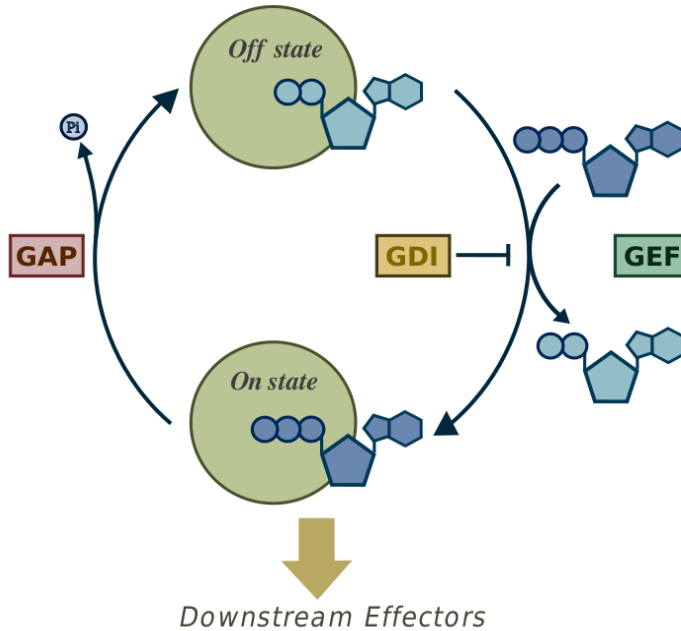
## 1.1 Guanine nucleotide-binding proteins (G proteins)

Biological cells use a specialized molecular machinery to process extracellular signals. Since the 1970s it was known that signal processing occurs in three steps: reception (at the membrane outer side), transduction (through the membrane towards the inner side), and amplification (by intracellular effectors), using terms borrowed from electronic engineering. It was also known that GTP-binding proteins were signal transducers, and in 1980 the first G protein was isolated from rabbit liver (Northub JK, 1980). The G protein consisted of three polypeptide chains of about 45, 35 and 10 kDa, named the  $G\alpha$ ,  $G\beta$ , and  $G\gamma$  subunits of the heterotrimeric G protein, the  $G\alpha$  subunit binding and hydrolyzing GTP. Since then, many types of G proteins have been discovered and characterized, differing in size, structure, oligomeric state and functional role.

G proteins are present in all three domains of life: Bacteria, Archaea, and Eukarya, where they mostly perform regulatory functions in diverse cellular processes, not only in cell signaling, underscoring their fundamental role (Wittinghofer & Vetter, 2011).

G proteins act like molecular switches. When bound to GDP, they are said to be inactive (off-state), with low affinity for effectors. When bound to GTP, they are said to be active (on-state), with high affinity for effectors. After GTP hydrolysis the protein returns to the GDP-bound off-state and the cycle starts again.

The affinity of G proteins is high for both GDP and GTP (in the pM-nM range in the presence of  $Mg^{2+}$ ), but the intracellular GTP concentration is approximately 10-fold larger than GDP (Zala et al., 2017), and GDP-bound G proteins spontaneously return to their GTP-bound on-state. However, the guanidine nucleotide dissociation rate is low, and nucleotide exchange is very slow. The rate of GTP hydrolysis is also low (on the order of  $10^{-5} s^{-1}$ ). The slow kinetics of these steps allow for the switching to be regulated by other proteins: Guanine nucleotide Exchange Factors (GEFs, which accelerate dissociation and, therefore, exchange), Guanine nucleotide Dissociation Inhibitors (GDIs, which decelerate dissociation), and GTPase-Activating Proteins (GAPs, which accelerate GTP hydrolysis by several orders of magnitude, to about  $1 s^{-1}$ ) (Wittinghofer & Vetter, 2011). GEFs promote the switch to the on-state, GDIs block that switch, and GAPs promote the switch to the off-state (**Figure 1**) (Geyer & Wittinghofer, 1997).



**Figure 1.** G proteins cycle between inactive GDP-bound and active GTP-bound states, which is regulated by guanine nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs), and guanine nucleotide dissociation inhibitors (GDIs).

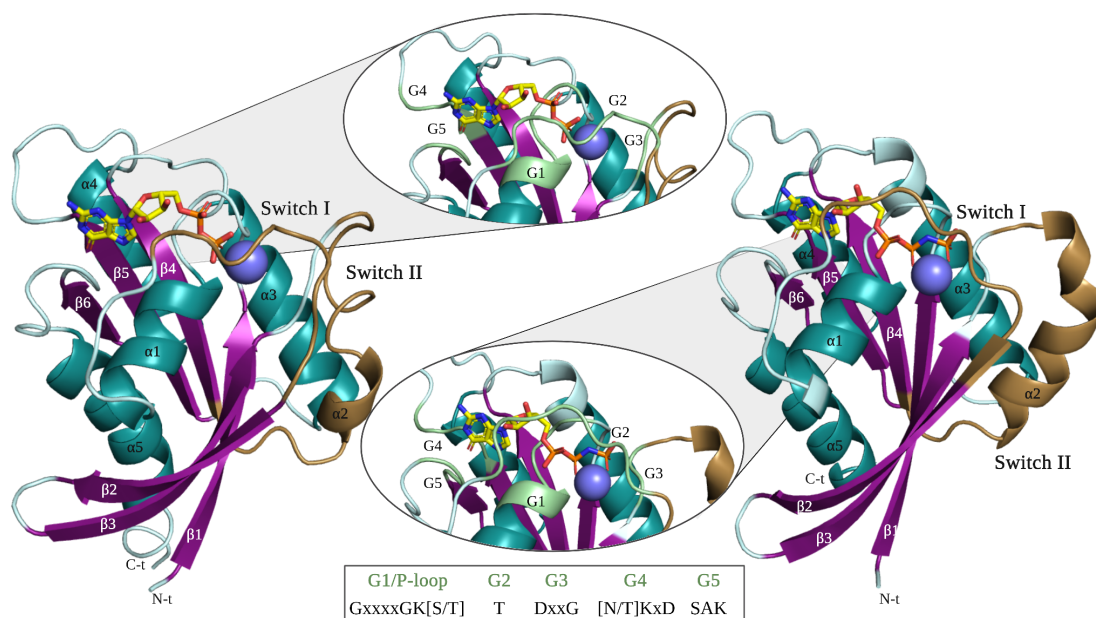
G proteins possess a common structural fold, known as the G domain (**Figure 2**), which binds and hydrolyzes GTP. The G domain contains five  $\alpha$ -helices, six  $\beta$ -strands, and multiple loops. Within these loops, there are five motifs that are essential for GTP binding and are conserved among all G proteins. Among these structural elements, two of them are known as switch I and switch II, whose conformation differs in the GDP- or GTP-bound states of the protein. The N-terminal and C-terminal regions of the G domain are variable in structure and are responsible for conferring specificity to different G protein families, as well as regulating their activity.

The conserved motifs within the loops are called G-boxes and numbered from G1 to G5. The G1 motif, also called the P-loop, has the consensus sequence of GxxxxGK[S/T]. G2, contained in the switch I region, is characterized by the presence of a highly conserved threonine residue. G3, contained in the switch II region, contains a conserved DxxG sequence. The G4 motif is characterized by the [N/T]KxD sequence. The sequence motif SAK is often associated with the G5 motif, but its conservation is not strong across all G proteins. Residues enclosed in brackets, indicate that both amino acids are frequently observed in that particular position,

while 'x' denotes any amino acid (Wittinghofer & Vetter, 2011). Boxes G1, G2 and G3 interact directly with the phosphates of the nucleotide, G1 predominantly with the  $\beta$ -phosphate, and G2 and G3 with the  $\gamma$ -phosphate. Boxes G4 and G5 interact with the base, and determine the specificity for guanosine nucleotide binding.

The switch regions derive their name from the conformational changes they undergo upon nucleotide exchange or GTP hydrolysis. This structural rearrangement enhances the affinity for effectors, playing a critical role in the activation and inactivation of the G protein signaling pathway, and mutations in these regions can contribute to various diseases and disorders (Wittinghofer & Vetter, 2011; Trabalzini & Retta, 2014; Orun, 2006).

The catalytic site is located in the pocket lined by switch I, switch II, and the P-loop. A glutamine residue in switch II and, in some G proteins, an arginine residue in switch I are essential for the catalytic activity. The  $\gamma$ -phosphate of GTP is positioned in the catalytic site by the P-loop and switch I.



**Figure 2.** Schematic representation of the structure of the H-Ras GTPase protein bound to GDP on the left (PDB: 1AA9) or bound to a non-hydrolysable GTP analog, GppNHp (guanosine 5'-[( $\beta$ , $\gamma$ )-imido]triphosphate), on the right (PDB: 1CTQ), where 18 and 23 residues from the C-terminus are not visible, respectively. The G-domain consists of five  $\alpha$  helices (turquoise) and six  $\beta$  strands (purple). The two switch regions (gold), which are involved in GTP binding and interactions with effectors, are also shown, along with the conserved G motifs that play a role in nucleotide binding. The magnesium ion is represented as a violet sphere.

G proteins require at least one Mg<sup>2+</sup> ion for proper function, which serves as a cofactor for nucleotide binding and hydrolysis. This cation stabilizes negative charges that arise during the transfer of the gamma phosphate from GTP to the acceptor molecule, such as water or another nucleophile, leading to the release of a phosphate ion and formation of GDP. Some G proteins require two Mg<sup>2+</sup> ions (Sprang, 2016). The activation of G proteins by GEFs is closely tied to the Mg<sup>2+</sup> ion, as they disrupt the magnesium binding site, thereby destabilizing the negative charges in the GDP phosphate group. This disruption causes the release of the phosphate groups first, followed by the release of the guanine base, facilitating nucleotide exchange (Bos et al., 2007). The concentration of total Mg<sup>2+</sup> inside mammalian cells is approximately 20 mM, most of it bound to different biomolecules, and about 1 mM is free in solution (Vink & Nechifor, 2011). High concentrations of Mg<sup>2+</sup> (typically 10 mM) are always used in assays with purified G proteins to saturate both of its states.

Attending to structural features most G proteins can be grouped into two families: small G proteins (of about 20 kDa, which are usually monomeric) and large heterotrimeric G proteins (with the GTP-binding  $\alpha$  subunit of about 45 kDa). Both contain the G domain and share the same switching mechanism. There are other GTP-binding proteins with different structural properties, like tubulin.

### 1.1.1 Small G proteins

Small G proteins consist of the G domain with N- and C-terminal regions of variable length. They are involved in intracellular cell signaling, not in extracellular signal transduction through the cytoplasmic membrane. But they require covalent lipidation to associate with cellular membranes. The most common lipid modifications are: prenylation (attaching a geranylgeranyl or farnesyl group to the C-terminal region), palmitoylation (attaching a palmitate group to one or more cysteine residues in the C-terminal region), and myristoylation (attachment of a myristate group at the N-terminal region of the protein).

The prototypic small GTPase proteins are those of the Ras family (from RAt Sarcoma). It was first described in the 1980s, and found to consist of three proteins in humans. Mutated Ras genes are common in cancer. Many of these mutations occur at the site of the glutamine residue in switch II (Q61 in H-Ras), yielding proteins unable to hydrolyze GTP and being permanently in the on-state.

The affinity of Ras proteins for GDP and GTP is similar (with a  $K_D$  in the pM range). They return to their GTP-bound on-state because of the 10-fold higher concentration of intracellular GTP (about 250  $\mu\text{M}$ ), but they do so very slowly, because the dissociation rate is low (about  $10^{-5} \text{ s}^{-1}$ ), and nucleotide exchange is very slow. GEFs reduce the affinity for GDP and can accelerate its dissociation several orders of magnitude. The rate of GTP hydrolysis is also low (on the order of  $10^{-4} \text{ s}^{-1}$ ), and can be accelerated by GAPs up to  $10 \text{ s}^{-1}$ .

The Ras superfamily includes other small G proteins that are classified based on their sequence homology, structural similarity, and functional properties into five different subfamilies: Ras, Rab, Rho, Ran and Arf. Each family has specific features and roles in different cellular processes (Trabalzini & Retta, 2014; Wennerberg et al., 2005).

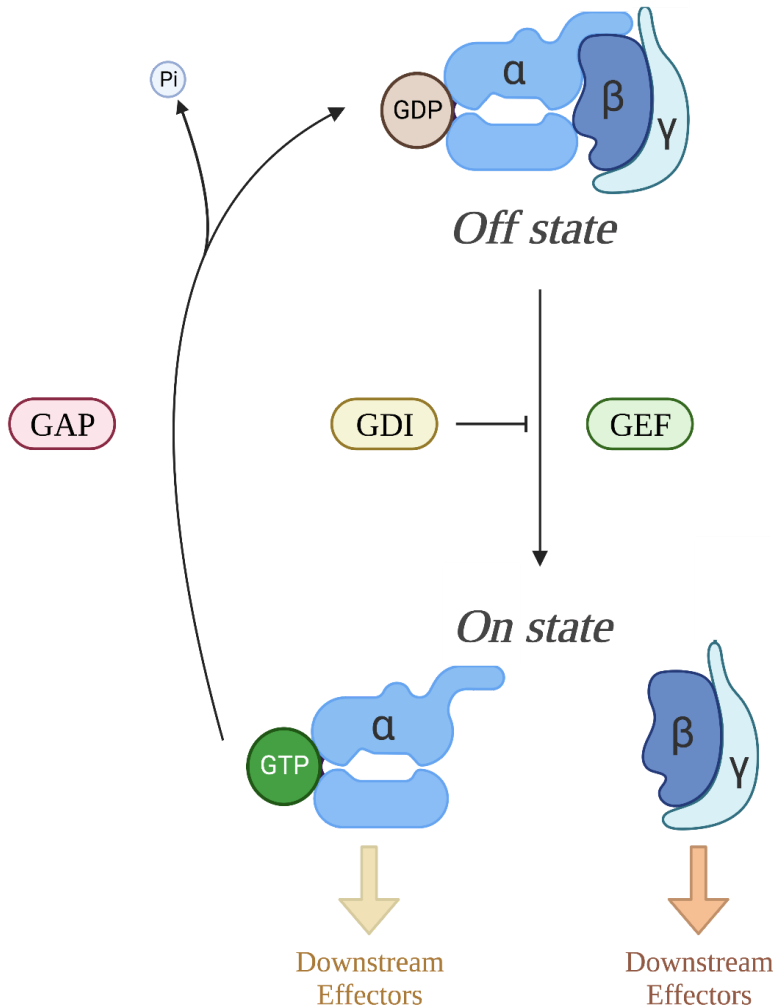
### **1.1.2 Large heterotrimeric G proteins**

Heterotrimeric G proteins are composed of three subunits, namely  $\alpha$ ,  $\beta$ , and  $\gamma$ , with molecular weights around 43 kDa, 35 kDa, and 10 kDa, respectively. This family of G proteins is primarily involved in signal transduction through the cytoplasmic membrane signaling processes. The  $\alpha$  and  $\gamma$  subunits undergo post-translational lipidation, palmitoylation and prenylation being most common. Palmitoylation involves attachment of the 16-carbon saturated fatty acid palmitate to a cysteine residue in the protein, commonly observed at the N-terminal region of  $G\alpha$  subunits. Prenylation attaches a 15- or 20-carbon isoprenoid to a cysteine residue near the C-terminus of the protein, commonly observed in  $G\gamma$  subunits. These modifications regulate the protein's membrane localization, stability, and interactions with other proteins (Wedegaertner et al., 1995).

The 40-45 kDa  $\alpha$  subunit of heterotrimeric G proteins ( $G\alpha$ ) consists of two structural domains: the N-terminal helical domain, and the C-terminal G domain. The helical domain confers specificity for interaction with other proteins, and the G domain binds GDP or GTP (and hydrolyzes GTP) suffering similar conformational changes as the small GTPases.

The  $\beta$  and  $\gamma$  subunits ( $G\beta\gamma$ ) form a tightly associated dimer that regulates the localization of the trimer, promoting its translocation to different subcellular compartments.

The regulatory cycle of the heterotrimeric G proteins is similar to that of the small G proteins. When bound to GTP,  $G\alpha$  is in the on-state and can interact with downstream effectors to generate a cellular response. Upon hydrolysis of GTP to GDP,  $G\alpha$  is in the off-state and dissociates from the effector proteins. However,  $G\alpha$  binding to GTP also causes its dissociation from  $G\beta\gamma$ . The  $G\beta\gamma$  dimer, which does not dissociate, can interact with downstream effectors independently of the alpha subunit (Syrovatkina et al., 2016) (**Figure 3**).

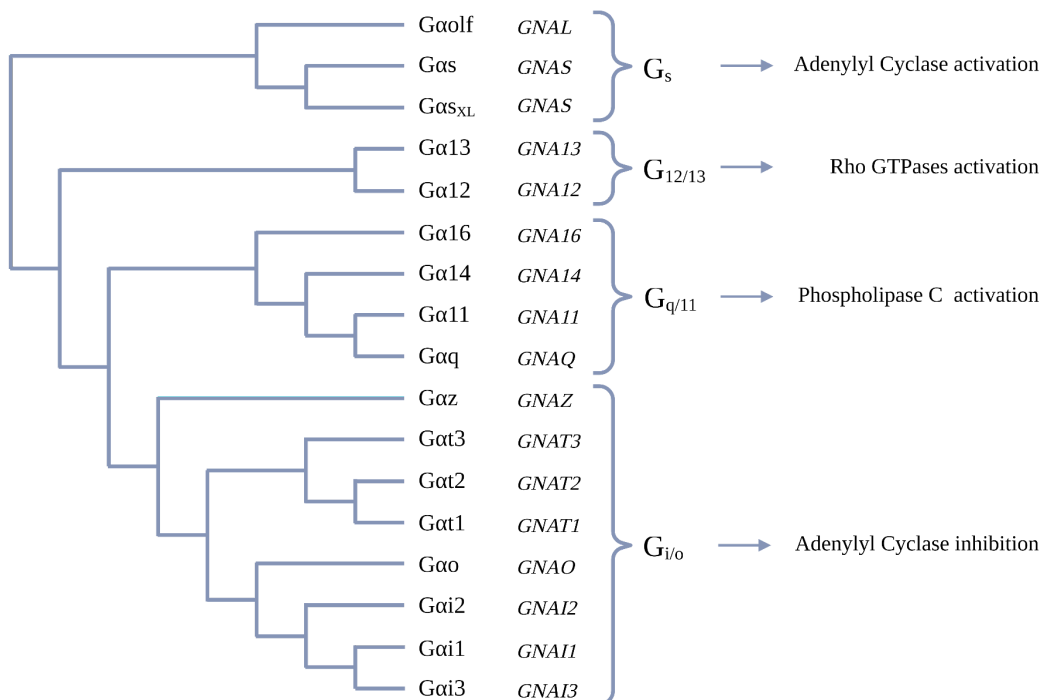


**Figure 3.** Activation cycle of heterotrimeric G proteins, which involves switching between inactive GDP-bound and active GTP-bound states. This process is regulated by GEFs, GAPs, and GDIs. Upon activation, the trimer dissociates into  $G\alpha$  and  $G\beta\gamma$  subunits, which subsequently interact with their corresponding effectors.



The affinity of  $G\alpha$  subunits is high for both GDP and GTP (in the nM range), and they return to their GTP-bound on-state because of the 10-fold higher concentration of intracellular GTP. But the guanine nucleotide dissociation rate is low (about  $10^{-3} \text{ s}^{-1}$ ), and exchange is very slow. The best characterized GEFs of  $G\alpha$  subunits are G Protein Coupled Receptors (GPCRs), transmembrane proteins that suffer conformational changes in response to extracellular signals and transduce them through the membrane by triggering  $G\alpha$  activation. They accelerate GDP dissociation by about two orders of magnitude. The rate of GTP hydrolysis is also low (approximately between  $10^{-1}$  and  $10^{-3} \text{ s}^{-1}$ ), and can be accelerated by GAPs by more than three orders of magnitude.

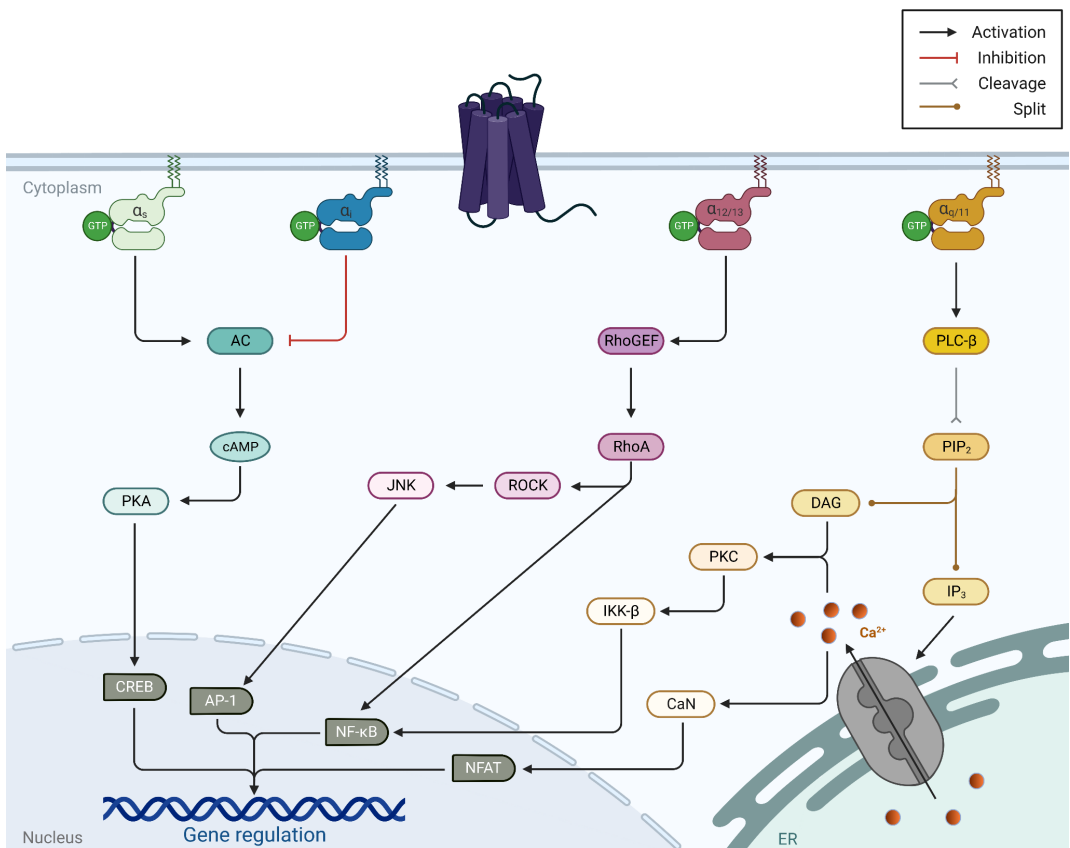
Heterotrimeric G proteins are classified into four families based on the signaling pathway regulated by their  $\alpha$  subunit:  $G_s$ ,  $G_{i/o}$ ,  $G_{q/11}$ , and  $G_{12/13}$ , as known in the early days of research on G proteins. However, new functions and signaling pathways have been discovered afterwards for many of these proteins, and the classification system is now considered to be somewhat outdated (**Figure 4**).



**Figure 4.** Guide tree and classification of  $G\alpha$  subunits and their corresponding signaling pathways. The genes that encode each subunit are also indicated. The subfamilies of heterotrimeric G proteins are shown, based on the signaling pathway regulated by their  $G\alpha$  subunit:  $G_s$ ,  $G_{i/o}$ ,  $G_{q/11}$ , and  $G_{12/13}$ . The tree was generated with Clustal Omega (McWilliam et al., 2013).

The number of G $\alpha$  present in humans is a topic of controversy among researchers (Jose & Choudhury, 2020; Wettschureck & Offermanns, 2005; Li et al., 2020). For the purposes of this introduction, we will refer to the Uniprot database, which includes 17 curated entries for canonical G $\alpha$  subunits in humans. Notably, the *GNAS* gene encodes two distinct isoforms, G $\alpha_s$  and G $\alpha_{sXL}$ , each with their own entry in Uniprot.

Human cells may produce five  $\beta$  (G $\beta$ 1-5) and twelve  $\gamma$  (G $\gamma$ 1-2, G $\gamma$ 2-5, G $\gamma$ 7-8, G $\gamma$ 10-13) proteins. Although there are 60 possible combinations to generate the G $\beta\gamma$  dimer and 1020 possible heterotrimer combinations, the actual number of combinations is smaller due to factors such as tissue-specific expression patterns, cellular localization, and trimer stability. For instance, G $\alpha$ 14 is mainly expressed in kidney, lung and spleen, while G $\beta$ 5 is primarily expressed in the brain, making unlikely their combination in a heterotrimer (Wettschureck & Offermanns, 2005).



**Figure 5.** Overview of the major signaling pathways activated by different heterotrimeric G protein families. The figure summarizes the most relevant downstream effectors and cellular responses mediated by the G<sub>s</sub>, G<sub>i/o</sub>, G<sub>q/11</sub>, and G<sub>12/13</sub> families of G proteins.

---

### 1.1.2.1 G<sub>s</sub> proteins

The 's' indicates its stimulatory effect on adenylate cyclase (AC), as the  $\alpha$  subunit of this subfamily stimulates AC leading to an increase in cytoplasmic cAMP concentration. This increase in cAMP levels activates protein kinase A, which in turn phosphorylates downstream targets, resulting in various physiological responses, including increased heart rate and glycogen breakdown (**Figure 5**). The G<sub>s</sub> protein subfamily comprises G<sub>s</sub>, its larger isoform G<sub>sXL</sub>, and G<sub>olf</sub>.

G<sub>s</sub> is the prototypical stimulatory G protein alpha subunit that is encoded by the *GNAS* gene and it is expressed in all human tissues (Wettschureck & Offermanns, 2005). G<sub>sXL</sub> is an isoform of G<sub>s</sub> that is only expressed in the neuroendocrine system (Wettschureck & Offermanns, 2005) and has a larger size due to the presence of a 682-residue long XL domain containing proline-rich motifs, which replaces the first 39 amino acids of G<sub>s</sub>. G<sub>sXL</sub>, but not G<sub>s</sub>, regulates clathrin-mediated endocytosis and plays a role in iron/transferrin uptake (He et al., 2017).

G<sub>olf</sub> is primarily expressed in the olfactory epithelium and several other neuroendocrine tissues, which is why it is referred to as 'olf' in its name, representing olfaction (Wettschureck & Offermanns, 2005). In olfactory receptor neurons, G<sub>olf</sub> couples to odorant receptors and stimulates adenylate cyclase, leading to the production of cAMP and activation of downstream signaling pathways (Goto, 2017).

### 1.1.2.2 G<sub>i</sub> proteins

The G<sub>i/o</sub> subfamily consists of G<sub>ai1</sub>, G<sub>ai2</sub>, G<sub>ai3</sub>, G<sub>at1</sub>, G<sub>at2</sub>, G<sub>at3</sub>, G<sub>ao</sub>, and G<sub>az</sub>.

G<sub>ai1</sub>, G<sub>ai2</sub>, and G<sub>ai3</sub> share approximately 80% sequence identity, and are extensively expressed throughout the human body (Wettschureck & Offermanns, 2005). The "i" in G<sub>ai</sub> stands for "inhibitory" because they primarily inhibit AC activity and reduce cAMP levels, which results in a decrease in PKA activity, in contrast to G<sub>s</sub>.

G<sub>ao</sub> is mainly present in neurons and neuroendocrine system (Wettschureck & Offermanns, 2005), and the "o" in its name refers to "other" because it can activate additional downstream effectors, including ion channels and protein kinases, in addition to inhibiting AC. G<sub>ao</sub> downstream effectors vary by cell type and extracellular signals (de Oliveira et al., 2019).

Gat1 and Gat2 are primarily found in rod and cone cells of the retina and are activated by light-sensitive GPCRs called rhodopsin and cone opsins. They regulate phototransduction and activate ion channels necessary for vision (Lerea et al., 1986). Gat3, also called gustducin or G<sub>α</sub>gust, is found in taste receptor cells and is activated by sweet and bitter taste receptors. It can activate PLC and increase intracellular calcium levels, causing taste neurotransmitter release. (Spielman, 1998).

Gaz plays a role in the modulation of neurotransmitter release, mainly inhibiting neurotransmitter release. Gaz is activated by GPCRs on the presynaptic neuron and can inhibit AC activity, leading to decreased cAMP levels and inhibition of neurotransmitter release from vesicles. Additionally, it is involved in the regulation of ion channels and intracellular calcium levels (Hinton et al., 1990).

#### 1.1.2.2.1 Gai3

Gai3 is a 354 amino acid protein expressed in various tissues and cell types. It preferentially interacts with Gβ1, Gβ2, Gβ4, Gγ1, Gγ5, Gγ7 and Gγ11 to form several G protein complexes (Tennakoon et al., 2021).

One of the most well-studied downstream effectors of Gai3 is adenylyl cyclase (AC), an enzyme that catalyzes the conversion of ATP to cyclic AMP (cAMP). Gai3 inhibits AC, leading to a decrease in cAMP levels and subsequent downstream signaling events. This has been linked to the regulation of cellular proliferation and differentiation, as well as the modulation of neurotransmitter release.

In addition to its role in modulating AC activity, Gai3 also interacts with a variety of other effector proteins. These include ion channels, such as the inward rectifying potassium channel (Ivanina et al., 2004), as well as phosphoinositide 3-kinase (PI3K). Gai3 also interacts with a number of downstream effectors, including protein kinase C (PKC) and mitogen-activated protein kinases (MAPKs) (Ghosh et al., 2008).

Gai3 has been implicated in a variety of physiological and pathological processes. In the nervous system, Gai3 has been shown to modulate neurotransmitter release and synaptic plasticity, and has been linked to the regulation of memory formation (Ivanina et al., 2004; Vellano et al., 2011). Gai3 has also been implicated in the regulation of cellular proliferation and differentiation, and has been linked to the development of cancer and other diseases (Ghosh et al., 2008).

### 1.1.2.3 G<sub>q/11</sub> proteins

The G<sub>q/11</sub> subfamily comprises Gα<sub>q</sub>, Gα<sub>11</sub>, Gα<sub>14</sub>, and Gα<sub>16</sub>, that, upon binding to GTP, activate PLC-β isoforms, leading to the generation of diacylglycerol (DAG) and IP<sub>3</sub>. DAG activates PKC, whereas IP<sub>3</sub> binds to IP<sub>3</sub> receptors on the endoplasmic reticulum, causing the release of Ca<sup>2+</sup> into the cytoplasm (**Figure 5**) (Kamato et al., 2015).

Gα<sub>q</sub> and Gα<sub>11</sub> are widely expressed in various tissues and organs throughout the human body and have been extensively studied. On the other hand, Gα<sub>14</sub> is primarily expressed in the kidney, lung, and spleen, while Gα<sub>16</sub> is mainly expressed in hematopoietic cells (Wettschureck & Offermanns, 2005). Gα<sub>14</sub> and Gα<sub>16</sub> share structural and functional similarities with Gα<sub>q</sub> and Gα<sub>11</sub> but are less well-understood. There is no Gα<sub>15</sub> in human cells. Gα<sub>15</sub> was described in murine cells and is the paralog of human Gα<sub>16</sub>.

Dysregulation of G<sub>q</sub> signaling has been linked to various disorders, including hypertension, heart failure, and cancer. Drugs that modulate G<sub>q</sub> signaling, such as angiotensin receptor blockers and calcium channel blockers, are widely used to treat these diseases (Kamato et al., 2017).

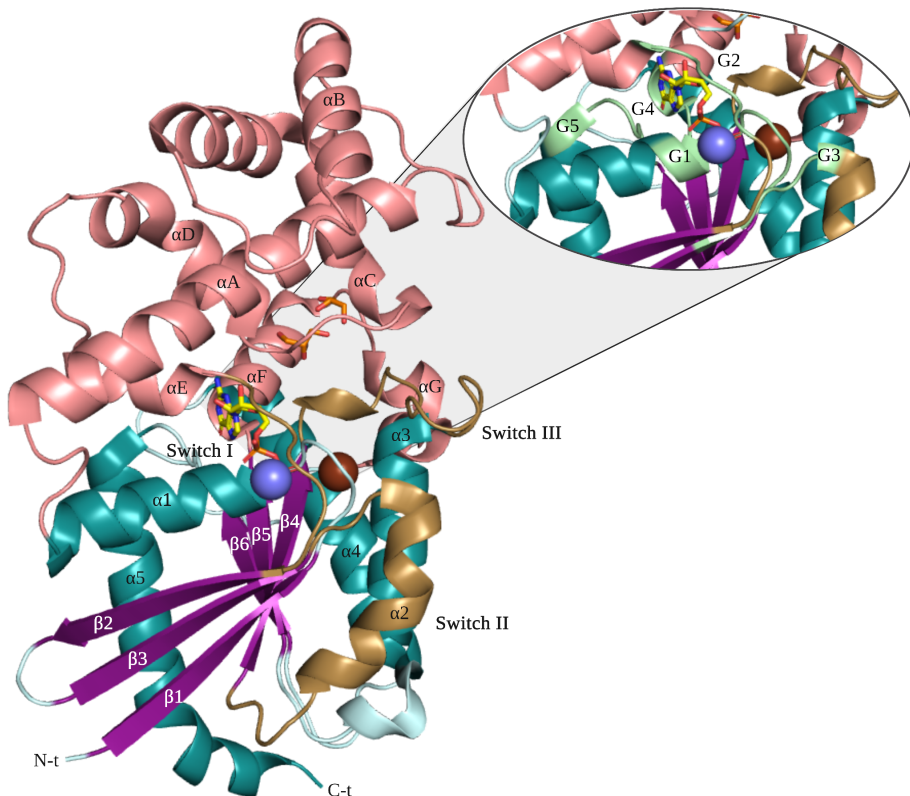
### 1.1.2.4 G<sub>12/13</sub> proteins

The G<sub>12/13</sub> subfamily consists of the closely related Gα<sub>12</sub> and Gα<sub>13</sub> proteins, with over 80% sequence identity. Upon activation, they activate downstream effectors such as RhoGEFs, which activate Rho family GTPases (RhoA, RhoB and RhoC). RhoA, in turn, activates downstream effectors such as Rho-associated protein kinase (ROCK) and mDia, which regulate various cellular processes including cytoskeletal reorganization, cell adhesion, migration, and proliferation (Guo et al., 2022). Gα<sub>13</sub> is implicated in regulation Hippo signaling and actin dynamics (Yagi et al., 2019).

The G<sub>12/13</sub> subfamily plays a role in various physiological processes, cancer cell invasion, and platelet activation. It is involved in signaling pathways related to cell growth, migration, and apoptosis. It promotes tumor cell invasion and mediates thrombin-induced platelet activation (Suzuki et al., 2009).

## 1.2 Structure of heterotrimeric G proteins

The  $\alpha$  subunit of heterotrimeric G proteins is made up of two domains: the G-domain and the helical domain (**Figure 6**). The G-domain is structurally similar to small GTPases such as Ras, thus, it is also called the Ras-like domain, and contains five  $\alpha$ -helices, six  $\beta$ -strands, and 10 loops. The helical domain contains 7 helices (named  $\alpha$ A-  $\alpha$ G), and is involved in the interaction with the  $\beta$  and  $\gamma$  subunits of the heterotrimeric G protein, as well as with effector proteins (Liu & Northup, 1998). The N-terminal disordered region (preceding the G-domain) and the switch II region are involved in interacting with G $\beta$  (see below). Apart from the two switch regions highlighted for monomeric G proteins, there is an additional switch III region, which does not directly interact with GTP, but undergoes a conformational change that stabilizes the new state of the G protein. In the G-domain of G $\alpha$  proteins, the  $\beta$ 5- $\alpha$ 4 loop is longer than in Ras proteins, including a short helical segment that is considered part of the helical domain ( $\alpha$ F).

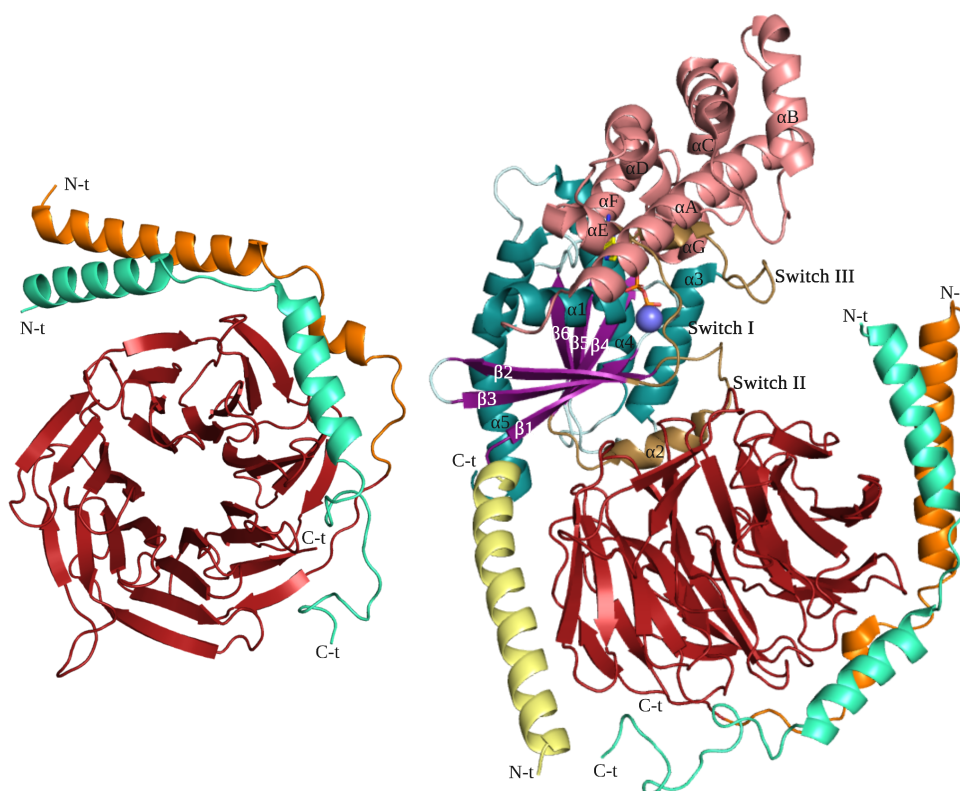


**Figure 6.** Ribbon representation of *Ga $\alpha$* -GDP crystal structure (PDB 6EG8, Chain J). The N-terminal 9 residues are absent. The helical domain is pink and the G-domain is turquoise (helices) and purple (strands). The three switch regions are in gold. The magnesium ion is represented as a violet sphere.

### 1.2.1 G $\beta$ $\gamma$ structure

The protein G $\beta$  has a molecular weight of 36-40 kDa. G $\beta$ 1-4 share over 80% sequence identity, while G $\beta$ 5 is more distinct, showing only 50% sequence similarity (Tennakoon et al., 2021). It consists of two distinct regions: an N-terminal alpha helix and a beta-propeller domain (**Figure 7**). The latter is made up of a repeating sequence known as a WD40 repeat. The  $\beta$ -propeller domain features seven antiparallel  $\beta$ -sheets arranged in a propeller shape. Each blade of the propeller comprises four twisted  $\beta$  strands (Pons et al., 2003).

G $\alpha$  binds to G $\beta$  $\gamma$  primarily through the interaction of the N-t region and switch II with blades 1, 2, and 3. The N-terminal region forms a long helix in the complex, and the switch-II region is located on the central tunnel of G $\beta$  (**Figure 7**) (Clapman, 1997).



**Figure 7.** On the left, the crystal structure of G $\beta$ 1 $\gamma$ 2 (PDB 5UKK) is depicted, with seven N-terminal and seven C-terminal residues from the G $\gamma$  structure absent. The  $\beta$ -propeller domain is depicted in red, the N-terminal helix in orange, and the  $\gamma$  subunit in green. On the right, the crystal structure of the GasG $\beta$ 1 $\gamma$ 2 heterotrimer (PDB 6EG8) is shown. The Gas subunit is represented in yellow (N-terminal helix), turquoise (helices), purple (strands) and gold (the three switch regions). The helical domain is shown in pink. The magnesium cation is represented as a violet sphere. The view of G $\beta$ 1 $\gamma$ 2 depicted in the left panel is rotated with respect to the view of the GasG $\beta$ 1 $\gamma$ 2 in the right panel.

G $\gamma$  has a molecular weight of 8-11 kDa and is composed of two  $\alpha$  helices that interact with different regions of the G $\beta$  propeller and with its N-terminal helix. Post-translational lipidation in the C-terminal region of this subunit facilitates its attachment to membranes. One interesting feature of G $\gamma$  is that it does not have intramolecular tertiary interactions. The association between G $\gamma$  and G $\beta$  is tight, and the two subunits do not dissociate upon activation of the G protein heterotrimer, in contrast to G $\alpha$  (Clapman, 1997; Tennakoon et al., 2021).

For structural studies the G proteins need to be produced in pure form. While some G $\alpha$  subunits can be synthesized in bacterial expression systems, G $\beta\gamma$  subunits are more challenging to synthesize. G $\beta\gamma$  can be formed in vitro in rabbit reticulocyte lysate, but only 30-50% of synthesized G $\beta$  can assemble into G $\beta\gamma$ . Chaperones are necessary for proper folding and assembly of G $\beta\gamma$ , which cannot be made natively in bacteria but can be produced in insect cells. Truncations of G $\beta$  can prevent correct assembly, whereas removal of certain amino acids from G $\gamma$  has a minor effect on assembly (Clapman, 1997).

### 1.3 Regulation of heterotrimeric G proteins

The regulation of heterotrimeric G proteins shares similarities with small G proteins, involving a variety of proteins that function as GEFs, GDIs, or GAPs mediating different aspects of their cycle.

#### 1.3.1 Guanine nucleotide exchange factors

GEFs catalyze the exchange of GDP for GTP in heterotrimeric G proteins in a similar way as occurs in small G proteins.

##### 1.3.1.1 G protein Coupled Receptors

The canonical GEFs of heterotrimeric G proteins are G protein-coupled receptors (GPCRs), the transducer of extracellular signals across the cytoplasmatic membrane.

GPCRs are very diverse, with over 800 different human GPCRs that can be categorized into six classes based on their structure and function. The Rhodopsin-like family (Class A) is the largest and most diverse and includes receptors for neurotransmitters, hormones, and sensory stimuli. The Secretin-like family (Class B) includes receptors for peptides while the Metabotropic glutamate family (Class C) includes receptors for the neurotransmitter glutamate. The Fungal mating pheromone



family (Class D) includes receptors for pheromones involved in fungal mating. The Cyclic AMP receptor protein family (Class E) includes receptors involved in bacterial chemotaxis and virulence. Finally, the Frizzled/smoothed family (Class F) includes receptors involved in the Wnt and Hedgehog signaling pathways.

The characteristic structure of G protein-coupled receptors (GPCRs) is a seven-transmembrane (7TM) alpha-helical bundle, with three extracellular and three intracellular loops. The N-terminus of the receptor is typically located outside the cell, while the C-terminus is located inside. GPCRs are most similar in their TM segments and have variable structures in their N- and C-terminal regions, and in the intracellular loop between TM5 and TM6 helices. The extracellular loops play a crucial role in ligand recognition, while the intracellular loops are involved in G protein coupling and signal transduction. GPCRs also contain conserved motifs such as the DRY motif and NPxxY motif that are important for receptor activation and signaling (Kobilka, 2007).

The approximately 800 GPCRs are very selective for their about 1,000 natural ligands. They are also selective for G $\alpha$  proteins, but some receptors can bind more than one of the 17 G $\alpha$  proteins, and different receptors can bind the same G $\alpha$  protein. Selectivity is based primarily on recognition between the extension of TM5 helix and  $\alpha$ 4,  $\beta$ 6 and  $\alpha$ 5 region opposite to the switch II helix (Flock et al., 2017). It is an open matter whether the receptor is associated (pre-coupling model) to the GDP-bound G protein before the signal arrives and activates it or not (random collision model). Understanding the mechanisms for coupling and selection is further compounded by the heteromerization of some GPCRs (Ferré, 2015). The prevailing model is that the affinity of the activated receptor for the G-protein is much higher in its nucleotide-depleted state, and this is the driving force for the conformational change that triggers the dissociation of the GDP (Garcia-Marcos, 2023). The binding of the abundant GTP reduces the affinity of the G $\alpha$  protein for the GPCR and for the G $\beta\gamma$  dimer causing dissociation from both and the subsequent activation of downstream signaling pathways.

GPCRs are considered one of the most significant membrane protein families responsible for critical cellular signaling and various physiological processes. Due to their importance, the pharmaceutical industry has identified GPCRs as major therapeutic drug targets. Currently, around 35% of marketed drugs are estimated to

target GPCRs, with most drugs targeting the Class A receptors. The drugs targeting GPCRs are used to treat a wide range of conditions, including allergies, hypertension, pain, depression, and cancer (Yang et al., 2021).

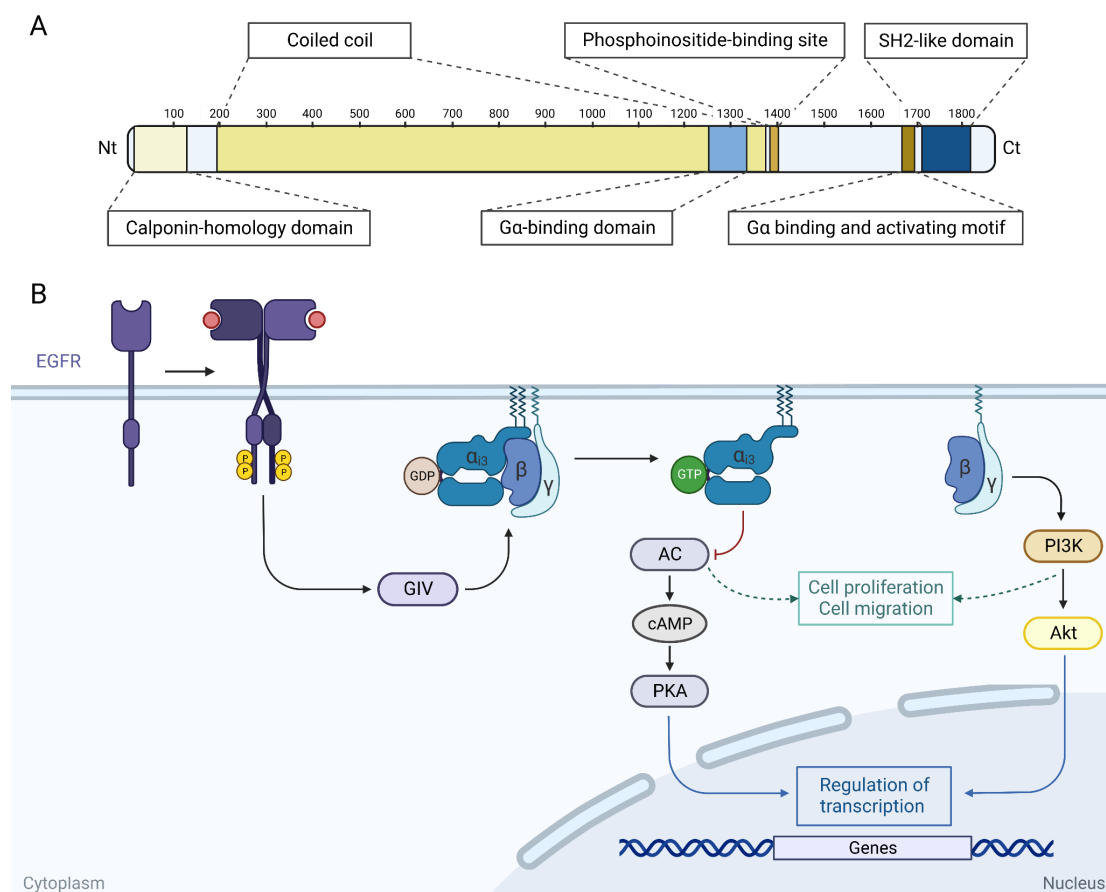
### 1.3.1.2 Non-receptor G protein GEFs

Intracellular GEFs can activate heterotrimeric G proteins in a receptor independent mode by accelerating GDP exchange for GTP between 2- and 10-fold, leading to the dissociation of the G $\alpha$  protein from the G $\beta\gamma$  dimer and triggering downstream signaling. Ric-8A, DAPLE and GIV are three such GEFs that activate Gai subunits. They contain a conserved G $\alpha$  binding and activating (GBA) motif responsible for binding to Gai. Ric-8A is involved in embryonic development and neuronal function, while DAPLE is known to modulate physiological processes such as planar cell polarity, cardiovascular development, and cancer metastasis (Srivastava & Artemyev, 2020; Aznar et al., 2015). GIV (G $\alpha$ -interacting vesicle-associated protein), also known as Girdin, is a multi-functional protein regulating cell proliferation, survival, adhesion, migration, and differentiation. The GBA motif consists of an approximately 20-residue-long sequence that is also known as the GEF motif (Garcia-Marcos et al., 2009). Because the same sequence of GIV that binds and activate Gai proteins (GEF activity), binds and inhibit the Gas proteins (GDI activity), it has been proposed to name it GEM (Guanine Exchange Modulator) (Ghosh et al., 2017). Inside cells the binding to Gai proteins occurs first (after EGF stimulation), and sequential binding to Gas occurs when S1674 is phosphorylated (Gupta et al., 2016).

### 1.3.1.3 GIV is a non-receptor GEF of Gai3

GIV is a large multifunctional protein of 216 kDa (**Figure 8A**), with a long region predicted to form coiled-coil structure and other regions predicted to be disordered. GIV contains two regions that can bind Gai proteins: the G-binding domain (GBD, residues 1343-1424) can bind both GDP-Gai and GTP-Gai proteins, and the GBA motif (1678-1694) preferentially binds GDP-Gai proteins. This region is thought to be flexible (Garcia-Marcos et al., 2009). The isolated C-terminal 1661-1871 fragment is indeed intrinsically disordered in solution (de Opakua et al., 2017). This fragment contains the GBA but not the GBD. It binds to Gai3 and increases nucleotide exchange in a dose dependent manner by a maximal factor of approximately 3, as measured by the increase in steady state GTP hydrolysis (Garcia-Marcos et al., 2009). This value is in the low end of the range measured for nonreceptor GEFs (between 2-

and 10-fold increase). Short fragments (between 19 and 35 residues) containing the GBA motif bind  $G_{\alpha i3}$  with  $K_D$  values between 0.6 and 1  $\mu\text{M}$  (DiGiacomo et al., 2017).



**Figure 8.** (A) Schematic representation of the different regions of GIV protein. (B) Scheme of signaling pathways initiated by EGFR stimulation by EGF and mediated by GIV. The GEF activity of GIV activates  $G_{\alpha i3}$  and dissociates  $G\beta\gamma$ , regulating proteins and genes relevant for cell proliferation and migration.

The N-terminal region of GIV binds to the intracellular domain of the Epidermal Growth Factor Receptor (EGFR). Upon binding of members of epidermal growth factor family (EGF) to the extracellular domain, the receptor undergoes a conformational change and becomes activated, resulting in the autophosphorylation of specific tyrosine residues in its intracellular domain, which then binds the C-terminal region of GIV. Binding of GIV to  $G_{\alpha i3}$  forms a ternary complex that links growth factor signaling to G protein-dependent pathways. AC inhibition reduces

cAMP levels and thus PKA activity and translocation to the nucleus, where it regulates transcription factors. PI3K activation leads to phosphorylation and translocation of Akt, regulating transcription factors and cell cycle regulators. (**Figure 8B**). The interplay of these and other pathways modulates the motogenic (migration) and mitogenic (proliferation) signals initiated by EGFR stimulation (Ghosh et al., 2010).

#### 1.3.1.4 Inhibitors of Gai3-GIV interaction and antimetastatic therapy

High levels of GIV expression are associated with poor prognosis and increased metastasis in various types of cancer. In rapidly growing, poorly motile breast and colon cancer cells and in noninvasive colorectal carcinomas in situ in which EGFR signaling favors mitosis over motility, a GEF-deficient splice variant of GIV was identified. In slow growing, highly motile cancer cells and late invasive carcinomas, GIV is highly expressed and has an intact GEF motif (Ghosh et al., 2010). The interaction between Gai3 and GIV has gained attention as a promising target in the fight against cancer due to its role in promoting metastasis. To identify small molecules capable of disrupting the interaction, an extensive screening process using both in silico and chemical high-throughput screening assays on the Library of Pharmacologically Active Compounds (LOPAC, with 1280 compounds, commercialized by Merck), identified two promising hits (DiGiacomo et al., 2017). One of them, named NF023, binds Gai3 and displaces the GBA peptide of GIV without interfering with Gai3-Gβγ binding. Although NF023 is not a valid lead (because it does not traverse the cytoplasmic membrane, and it also binds other proteins), it confirms the potential for targeting the Gai3-GIV interaction with drugs.

Based on these previous findings, our collaborators conducted a small-molecule screening for inhibitors of the Gai3-GIV interaction using 200,000 compounds. Among them, 580 hits were identified through fluorescence polarization assays, and 155 of these hits were further confirmed using an orthogonal secondary assay. After eliminating compounds with unfavorable chemical properties and availability issues, the remaining 69 compounds were designated as IGGi (Inhibitors of the GIV-Gai3 interaction). To evaluate their efficacy, these compounds were tested in MDA-MB-231 cells to determine their ability to inhibit cell migration without impacting cell viability. Subsequently, 44 compounds that demonstrated the desired effects were subjected to GST-fusion pull-down assays involving GIV and Gai3, with

---

NF023 serving as a control. Remarkably, only one compound, IGGi-11, exhibited the capability to inhibit G $\alpha$ i3 binding to GIV in this assay, making it a potential lead compound for therapeutic applications (Zhao et al., 2023).

### 1.3.2 GDIs and GAPs

Contrary to GEFs, these regulatory proteins deactivate G $\alpha$  subunit signaling by stabilizing the GDP-bound off-state or accelerating GTP hydrolysis. They commonly contain one of two motifs: the GoLoco motif (G $\alpha$ i/o-Loco) or the Regulators of G Protein Signaling (RGS) motif.

The GoLoco motif (also known as G protein regulatory, GPR, motif) is ~20 amino acids in length (Siderovski et al., 1999) and is present in proteins which serve as GDI for G $\alpha$ i proteins. The GoLoco motif binds to GDP-bound G $\alpha$ i subunits with nM affinity stabilizing them in the off-state (Kimple et al., 2001). The crystal structure of human GDP-G $\alpha$ i1 bound to a 36-residue-long fragment of RGS14 (Kimple et al., 2002) shows the peptide interacting with the G-domain through its N-terminal GoLoco sequence (folded into a helix and inserted in the cleft between switch II and  $\alpha$ 3 of G $\alpha$ i1). The C-terminal non-conserved region of the peptide interacts with the helical domain, providing specificity for G $\alpha$ i proteins. The conformational changes in the switch regions and a direct interaction with the phosphates of the GDP are thought to cause the GDI activity of the peptide (**Figure 9**).

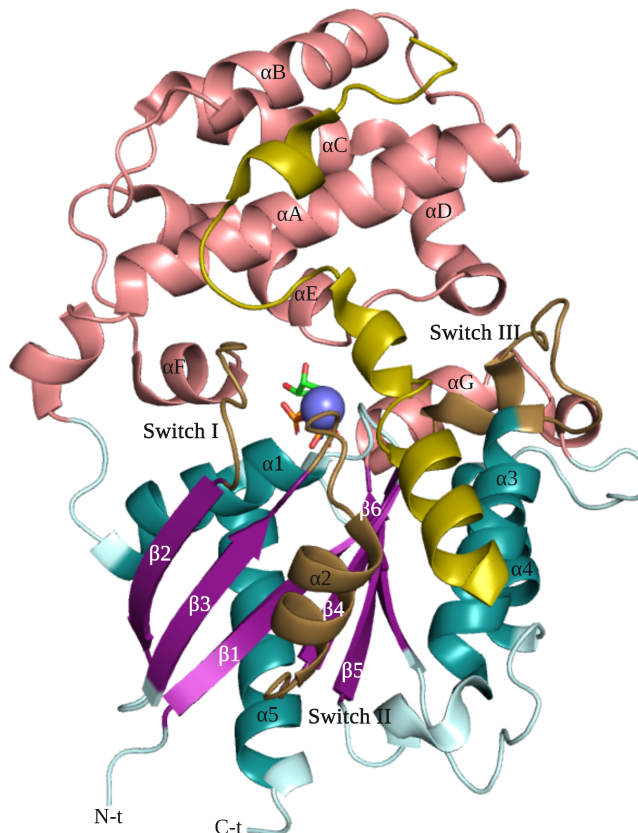
The RGS motif is a conserved domain of about 120 amino acids folded into a bundle of helices. It is present in a variety of proteins, including RGS ones. These proteins act as GTPase-activating proteins (GAPs) for G $\alpha$ i/o, G $\alpha$ q/11, and G $\alpha$ 2/13 proteins, accelerating the hydrolysis of GTP to generate GDP-bound off-state G protein. The action of RGS proteins results in a reduction in the magnitude and duration of downstream signaling events, allowing for the proper balance of signaling (Siderovski & Willard, 2005). Co-crystal structures of RGS domains with GDP-AlF<sub>4</sub>-G $\alpha$  proteins show the RGS bound to a cleft lined by switch III and the middle region of the connection  $\beta$ 5- $\alpha$ 4. The tetrafluoroaluminate ion occupies the position corresponding to the  $\gamma$ -phosphate of GTP-bound G $\alpha$ , creating an analog of the transition state of GTP hydrolysis (Coleman et al., 1994). RGS activates GTP hydrolysis by stabilizing the transition state.

### 1.3.2.1 Regulator of G protein Signaling 12

RGS12 is a large protein of 156 kDa containing several functional domains, including PDZ, phosphotyrosine binding, Ras binding, and RGS domains.

The RGS domain, spanning residues 715-832, classifies RGS12 as a member of the RGS family. This sequence specifically contributes to the protein's GAP activity, enabling it to reduce Gai/o protein signaling by accelerating its GTP hydrolysis (Siderovski & Willard, 2005).

The GoLoco motif, located between residues 1187-1209 allows RGS12 to act as GDI, stabilizing Gai proteins in their GDP-bound off-state. The RGS12<sub>1185-1221</sub> fragment binds human GDP-Gai3 with a  $K_D$  of  $54 \pm 4$  nM at room temperature.



**Figure 9.** Crystal structure (PDB 1KJY) of human GDP-Gai1 bound to rat RGS14<sub>496-531</sub> (yellow). The helical domain is pink and the G-domain is turquoise (helices) and purple (strands). The three switch regions are in gold. The magnesium ion is represented as a violet sphere.



structure the peptide is folded into a  $\beta$ -hairpin with its C-terminal strand inserted into the hydrophobic cleft between switch II and  $\alpha 3$ . KB-1753 competes with RGS4 for binding to GDP-AlF<sub>4</sub>-Gα<sub>13</sub>, suggesting that it can be used to reduce the GAP activity of RGS proteins (Johnston et al., 2008).





# Hypothesis and Research Objectives



The protein-protein interaction Gai3-GIV is dysregulated in several types of cancer. The interaction promotes cell migration, and high levels of GIV correlate with tumor invasiveness. Disrupting the Gai3-GIV interaction might help avoid metastasis in human cancers. The interaction has recently been structurally characterized by NMR and molecular modeling, and the binding interface has been confirmed as druggable. Shortly before the initiation of this thesis, a screen of a large library of small molecules (200,000 organic compounds) yielded several inhibitors of the interaction, and one of them (named IGGi-11) was found to inhibit pro-invasive traits of metastatic breast cancer cells without interfering with the canonical regulation of Gai3 by GPCRs.

### **Hypothesis:**

IGGi-11 disrupts the interaction between human Gai3 and GIV in a competitive manner by binding to the same site on Gai3, and a high resolution structure of the complex will provide details on the molecular recognition useful to improve its affinity.

### **Objectives:**

#### **2.1 Characterize the binding of IGGi-11 to human Gai3**

The aim is to identify the binding site, measure the affinity of the interaction, and compare with the binding of GIV.

#### **2.2 Obtain the crystal structures of Gai3 bound to GIV and bound to IGGi-11**

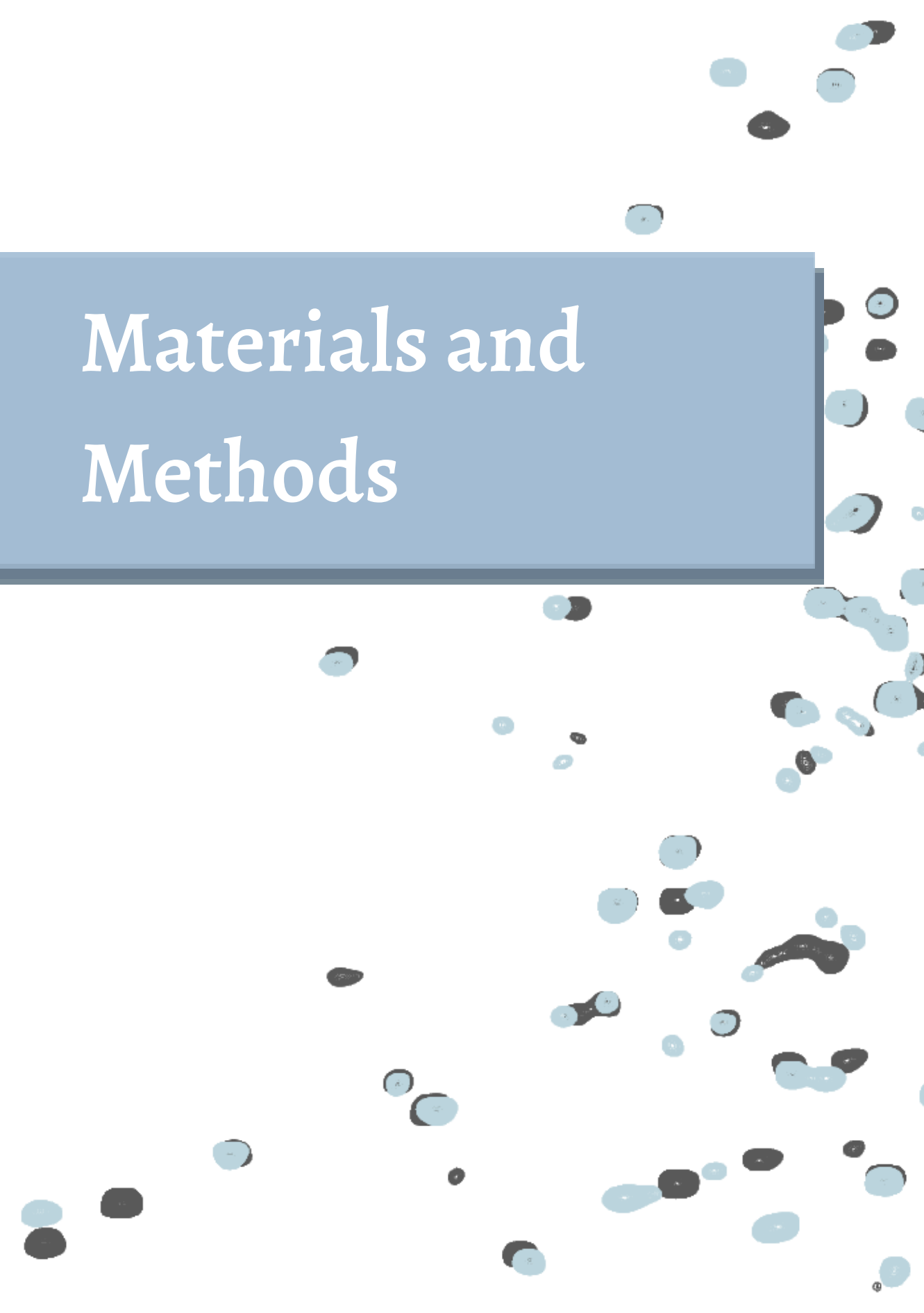
The aim is to obtain high resolution structures of the complexes to understand the bases of the molecular recognition.

#### **2.3 Characterize the binding of Gβ1γ2 and other Gai3 regulators: RGS12 protein and the non-natural peptides R6A and KB-1753**

The aim is to identify the binding sites and compare them with GIV's binding site.



# Materials and Methods





### 3.1 Protein expression and purification

#### 3.1.1 Gene design for human Gai3 protein production in bacterial cells

The clone for production of human full-length Gai3 (Uniprot entry P08754, isoform 1) was provided by professor I. Shimada (Mase et al., 2012). The gene (cloned in plasmid pET24d(+) under the control of T7 promoter) coded for an N-terminal His<sub>10</sub> tag (for protein purification by immobilized metal affinity chromatography, IMAC), followed by the HRV-3C protease cleavage sequence (to remove the tag) and four non-native residues preceding the initial methionine of Gai3. The Gai3 W258F clone was obtained by site directed mutagenesis.

The gene used to produce the shortened human Gai3<sub>31-350</sub> (sGai3) protein was designed with an N-terminal fusion to ubiquitin to favor gene expression and protein solubility in bacterial cultures (Rogov et al., 2012). The ubiquitin moiety contains a His<sub>10</sub>-insertion in the loop between the first  $\beta$ -hairpin and the helix, and is followed by the TEV protease cleavage sequence before the native A31 residue. The gene, with optimized codons for *E. coli*, was purchased from Genscript cloned in plasmid pET29a(+) under the control of T7 promoter.

**Table 1.** Sequences of Gai3 and sGai3 proteins used in this project. Protease cleavage sites are indicated in blue and non-native residues remaining after cleavage are underlined. The residue in red was mutated to phenylalanine to obtain Gai3 W258F.

Protein	Sequence
Gai3 (1-354)	MGHHHHHHHHSSGHPDDE <u>VLFQGPDDH</u> MGCTLSAEDKAAVERSKMIDRNLREDGE KAAKEVKLLLLGAGESGKSTIVKQMKIIHEDGYSEDECKQYKVVVYSNTIQSIIAII RAMGRLKIDFGAARADDARQLFVLGSAEEGVMTPELAGVIKRLWRDGGVQACFSR SREYQLNDSASYLLNDLDRISQSNIPTQQDVLRLTRVKTGTGIVETHFTFKDLYFKMF DVGGQRSEKRWIHCFFEGVTAIIFCVALSVDYDLVLAEDEEMNRMHESMKLFDSICNN <u>KW</u> FTETSIIILFLNKKDLFEKIKRSPLTICYPEYTGSNTEYEEAAAYIQCFEDLNRR KDTKEIYTHFTCATDTKNVQFVFDVAITDVIKNNLKECGLY
sGai3 (31-350)	MQIFVKTLTGKTTITLEVEPGSAHHHHHHHHHAGSSDTIENVKAKIQDKEGIPPDQQ RLIFAGKQLEDGRTLSDYNIQEKSTLELVLELQNESGGENLYFQGAKEVKLLLLGAG ESGKSTIVKQMKIIHEDGYSEDECKQYKVVVYSNTIQSIIAII RAMGRLKIDFGAAR RADDARQLFVLGSAEEGVMTPELAGVIKRLWRDGGVQACFSRSREYQLNDSASYLL NDLDRISQSNIPTQQDVLRLTRVKTGTGIVETHFTFKDLYFKMFDVGGQRSEKRWIHC CFEGVTAIIFCVALSVDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTETSIIILFLNK KDLFEKIKRSPLTICYPEYTGSNTEYEEAAAYIQCFEDLNRRKDTKEIYTHFTCAT DTKNVQFVFDVAITDVIKNNLKE

### 3.1.2 Protein production and purification

The protocol for protein production and purification is summarized in **Figure 11**.

#### 3.1.2.1 Gai3 production

Because some of the human codons in the human Gai3 gene are rare in bacteria, *E. coli* BL21(DE3) Rosetta cells (harboring a plasmid coding for tRNAs corresponding to codons that are rare in bacterial but not in human genes) were employed for its expression. In all growth media used for Gai3, it was necessary to include antibiotics (30 mg/l kanamycin and 34 mg/l chloramphenicol) for the selection of cells containing both the plasmid carrying the gene for Gai3, which is resistant to kanamycin, and the chloramphenicol-resistant plasmid present in Rosetta cells.

For natural isotopic abundance, cells were cultured in auto-induction ZYP-5052 medium (Studier, 2005). The medium was inoculated with transformed cells from a saturated preculture (grown in peptone-based medium) to an initial OD<sub>600</sub> = 0.05, and the cells were grown at 37 °C for 2 hours. The temperature was lowered to 20 °C to promote continued growth and protein production for 16-20 hours. The yield was about 11 mg of pure protein per liter of culture.

To achieve uniform <sup>2</sup>H, <sup>15</sup>N isotopic enrichment, the cells were cultured in a modified M9 minimal medium containing 1 g/L <sup>15</sup>N<sup>2</sup>H<sub>4</sub>Cl, 2 g/L <sup>2</sup>H-glucose, and 1 g/L of <sup>2</sup>H-<sup>15</sup>N Celtone base powder in <sup>2</sup>H<sub>2</sub>O. The cells were grown at 37 °C until they reached an OD<sub>600</sub> = 0.75. Protein expression was then induced by adding 1 mM IPTG to the cell cultures, followed by incubation for 16 hours at 20 or 23 °C. The yield was about 8 mg of pure protein per liter of culture.

For selective <sup>15</sup>N isotope enrichment of Trp residues, the cells were initially cultured at 37 °C in a modified M9 minimal medium without isotopic enrichment. Once the cells reached an OD<sub>600</sub> of 1.4, they were centrifuged, and the pellet was resuspended in prewarmed M9 minimal medium containing 100 mg/l of <sup>15</sup>N-Trp and 100 mg/l of the other 19 amino acids with natural isotopic abundance. The cells were then grown for 15 minutes before induction of protein production with 1 mM IPTG for 16-20 hours at 20 °C. A yield of 8.9 mg of Gai3 W258F and 3.9 mg of Gai3 per liter of culture was obtained.



### 3.1.2.2 sGai3 production

*E. coli* BL21(DE3) cells harboring the sGai3 plasmid were grown in peptone-yeast extract-NaCl medium with 30 mg/l kanamycin. Cells were grown at 37 °C until  $OD_{600} = 0.8$ , when expression was induced with 1 mM IPTG for 16–20 h at 20 °C. The yield was about 5 mg of pure protein per liter of culture.

### 3.1.2.3 Protein purification

Cells were harvested by centrifugation, resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 100 mM KCl, 1 mM DTT with EDTA-free protease inhibitors, flash-frozen in liquid nitrogen and stored at -80 °C. Thawed cells were sonicated and ultracentrifuged. All forms of Gai3 were predominantly found in the soluble fraction. Protein production in the cultures and its presence in the soluble or insoluble fractions was evaluated by reducing SDS-PAGE in 12 % acrylamide/bisacrylamide (29:1, by weight) gels run with Tris/Gly buffer and stained with Coomassie brilliant blue.

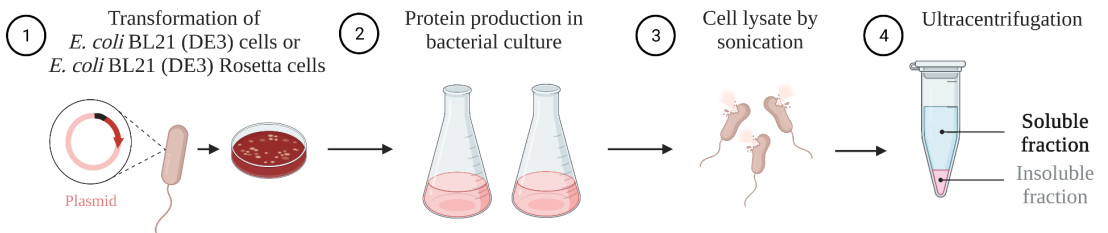
The proteins were purified by IMAC on 5 ml His-Trap columns loaded with  $Ni^{2+}$  ions and equilibrated in different buffers. For Gai3 and Gai3 W258F mutant, the buffer contained 50 mM Tris pH 8.0, 200 mM NaCl, 100 mM KCl, 20  $\mu$ M GDP, and 1 mM DTT. For sGai3, the buffer contained 50 mM Tris pH 8.0, 300 mM NaCl, 5% glycerol, 20  $\mu$ M GDP, and 1 mM DTT. After loading the sample and washing the column with buffer containing 25 mM imidazole, the protein was eluted using buffer containing 500 mM imidazole. Subsequently, dialysis was performed at 4 °C against buffer without imidazole for a duration of 15-20 hours, at the same time that the protein was cleaved by the addition of His-tagged HRV-3C protease (Gai3 and Gai3 W258F) or His-tagged TEV protease (sGai3). The ratio used was 1 mg of protease per 30 mg of total protein (estimated by absorbance at 280 nm).

The samples were loaded on an equilibrated His-Trap column to collect the cleaved protein in the flow-through. The flow-through was concentrated and separated by size exclusion chromatography (SEC) on a Superdex 200 column. The SEC was carried out using different buffers depending on the protein variant. For Gai3 and Gai3 W258F, the buffer consisted of 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 100 mM KCl, 20  $\mu$ M GDP, and 1 mM DTT. In the case of sGai3, the buffer composition was 50 mM Tris pH 8, 300 mM NaCl, 5% glycerol, 20  $\mu$ M GDP, and 1 mM DTT.

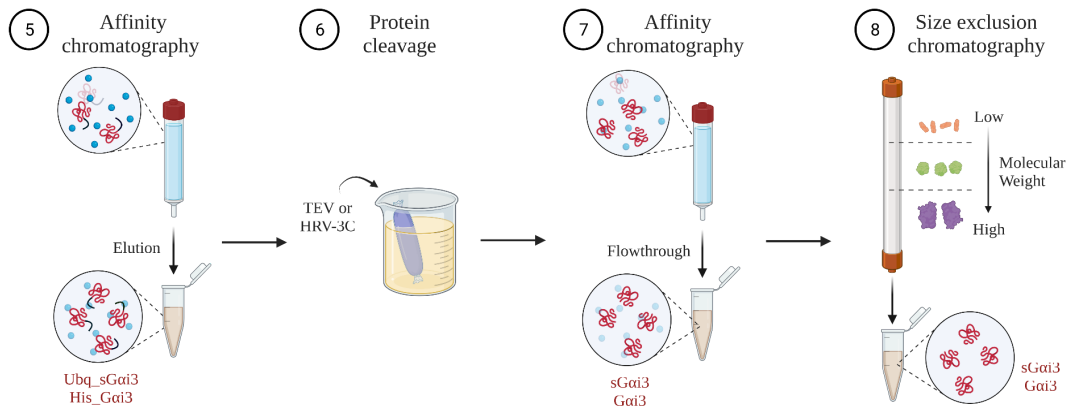
The purity, concentration, and identity of the purified proteins were confirmed by Coomassie-stained SDS-PAGE, ultraviolet light absorbance spectroscopy, and MALDI-TOF mass spectrometry.

Pure proteins were concentrated by ultrafiltration and the concentration was measured by absorbance at 280 nm using the extinction coefficient calculated from the amino acid sequence on the ExPASy ProtParam web server (Gasteiger et al., 2003) adding the contribution of the bound guanine nucleotide (Smith & Rittinger, 2002). Aliquots of the pure proteins were frozen in liquid nitrogen and stored at -80 °C until used. When necessary, the aliquots were thawed on ice and buffer exchanged by three cycles of dilution-concentration using centrifugal filters with 10 kDa cut-off membranes.

#### Growth



#### Purification



**Figure 11.** Summary of the protocol for protein production in bacterial cultures and purification by chromatography. The harvested cells were lysed and the proteins were purified from the soluble fraction by IMAC on His-Trap columns, followed by dialysis and tag elimination by proteolysis. Further purification was performed using size exclusion chromatography.

### 3.2 Ligand sample preparation for binding studies

A summary of the Gai3 ligands studied in this thesis is in Table 2.

**Table 2.** Properties of Gai3 ligands studied in this thesis. Uniprot entries for proteins are in italics. Protease cleavage sites are indicated in blue, native residues are in bold and non-native residues remaining after cleavage are underlined.

Protein	Subunit/ Fragment	Sequence
Gβ1γ2	β1 <sup>a</sup> <i>P62873</i>	HHHHHHHHSELDQLRQEAEQLKNQIRDARKACADATLSQITNNI DPVGRIQMRTRRTLGRHGLAKIYAMHWGTDLSRLVSVASQDGKLI WDSYTTNKVHAIPLRSSWVMTCAYPASGNVACGGLDNICSIYN LKTREGNVRVRSRELAGHTGYLSCCRFLDDNQIVTSSGDTTCALW DIETGQQTTFTGHTGDVMSLSLAPDTRLFVSGACDASAKLWDV REGMCRQTFTEGHESDINAICFFPNGNAFATGSDDATCRLFDLRA DQELMTYSHDNIICGITSVSFSSKSGRLLLAGYDDFNCNVWDALK ADRAGVLAGHDNRVSCLGVTDDGMAVATGSWDSFLKIWN
	γ2 <sup>b</sup> , <i>P59768</i>	MASNNTASIAQARKLVEQLKMEANIDRIKVSAAAADLMAYCEAH AKEDPLLLTPVPAENPFREKKFFSAIL
GIV <i>Q3V6T2-3<sup>c</sup></i>	GIV <sub>1660-1870</sub> (GIV <sub>C-t</sub> )	MGSSHHHHHHSSGLVPRGSHMSETLESRHHKIKTGSPGSEVVTL QQFLEESNKLTSVQIKSSSQENLLDEVMSLSVSSDFLGKDKPV SCGLARSVSGKTPGDFYDRRTTKPEFLRPGPRKTEDTYFISSAG KPTPGTQGKIKLVKESLSRQSKDSNPYATLPRASSVISTAEGT TRRTSIHDFLTKDSRLPISVDSPPAAADSNTTAASNVDKQVESR NSKRSRSREQQSS
	GIV <sub>1671-1705</sub>	KTGSPGSEVVTLQQFLEESNKLTSVQIKSSSQENL
	GIV <sub>1671-1692-DY<sup>d</sup></sub>	KTGSPGSEVVTLQQFLEESNKLDY
RGS12 <i>O14924</i>	RGS12 <sub>1185-1221</sub>	DEAEFFELISKAAQSNRADDQRGLLRKEDLVLPFLR
Non-natural peptide		Sequence
	R6A	MSQTKRLDDQLYWWEYL
	R6A-1	DQLYWWEYL
	KB-1753	SSRGYYHGIWVGEEGRLSR
Compound	IUPAQ name	
IGGi-11	4,4'-((9H-fluorene-2,7-disulfonyl)bis(methylazanediy))dibutyric acid	
IGGi-11me	Dimethyl 4,4'-((9H-fluorene-2,7-disulfonyl)bis(methylazanediy))dibutyrate	

<sup>a</sup> The initial methionine of Gβ1 after the His-tag is absent.

<sup>b</sup> Native C68 was mutated to serine to prevent post-translational geranylgeranyl modification.

<sup>c</sup> This isoform has a one-residue deletion as compared with the canonical Q3V6T2 one. We use this shorter isoform for residue numeration for consistency with many previous publications.

<sup>d</sup> Non-native DY sequence added at the C-terminus.

### 3.2.1 Protein fragments and non-natural sequences

GIV<sub>C-t</sub> was prepared as previously described (de Opakua et al., 2017). Briefly, a pET28b(+) plasmid containing a gene coding for His-tagged human GIV residues 1661–1871 was transformed in *E. coli* BL21(DE3). Cells were grown in autoinduction ZYP-5052 medium with 30 mg/l kanamycin for 16 h at 20 °C. The soluble protein was purified by IMAC and gel filtration, without proteolytic removal of the N-terminal tag. The yield was 0.2 mg of pure protein per liter of culture.

The other protein fragments or non-natural sequences were obtained as synthetic peptides used for NMR assays were obtained in powder form (purified by reverse phase chromatography and lyophilized) from collaborators or commercial suppliers. For NMR, crystallography, or calorimetry experiments they were prepared as concentrated solution stocks in different buffers depending on their solubility.

Peptide GIV<sub>1671-1705</sub> was solubilized in deuterated DMSO, and the concentration was calculated by gravimetry (5.3 mM).

Peptide GIV<sub>1672-1693-DY</sub> was designed based on the crystal structure of the rat G $\alpha$ i3 homolog bound to GIV<sub>1672-1693</sub> (Kalogriopoulos et al., 2019). The GIV residues modeled in the crystal (1674-1691) plus two at each end were considered those contributing to binding. The sequence DY was added at the C-terminus to increase solubility and measure concentration by absorbance. The powder was highly soluble in water, and less soluble after pH was adjusted to 7.0 (2.9 mM, measured by absorbance at 280 nm). For calorimetry a 1 mM stock was prepared in NMR buffer without GDP by dialysis. For crystallography, the powder was dissolved in 25 mM Tris pH 8.0, 150 mM NaCl and the concentration was measured by light absorbance at 280 nm (4 mM). The pH was adjusted to 8.0 using concentrated NaOH and HCl.

Peptide RGS12<sub>1185-1220</sub> was dissolved in water, the pH adjusted to 7.0, lyophilized and redissolved in NMR buffer without GDP. Peptide concentration (2.6 mM) was determined by absorbance at 205 nm (Anthis & Clore, 2013).

Peptide R6A was dissolved in deuterated DMSO, then diluted 2-fold with NMR buffer (10 mM HEPES pH 7.0, 10 mM MgCl<sub>2</sub>, 5 mM DTT, and 300  $\mu$ M GDP), the pH adjusted to pH 7.0 with concentrated NaOH and HCl, and the concentration measured by light absorbance at 280 nm (4.3 mM).

Peptide R6A-1 was dissolved in water with NaOH (at pH 12.3), the pH adjusted to 7.0 with HCl. Concentration was measured by light absorbance at 280 nm (1 mM).

Peptide KB-1753 (Johnston et al., 2006) was dissolved in NMR buffer plus 33% (by volume) deuterated DMSO. The pH was measured and not adjusted (pH = 6.6). Concentration was measured by light absorbance at 280 nm (1.9 mM).

### 3.2.2 Chemical compound

IGGi-11, in powder form, was purchased from Chembridge or from Ambinter. A 100 mM solution stock was prepared by dissolving the powder in deuterated DMSO. The concentration was not measured, it was calculated from the weight of the powder. For NMR experiments this stock, or a 2-fold dilution in NMR buffer, was used. For ITC it was diluted to 1 mM in 10 mM HEPES, 10 mM MgCl<sub>2</sub>, 1 mM TCEP, 300 μM GDP, pH 7.0. For crystallography IGGi-11 was diluted in 25 mM Tris pH 8.0, 150 mM NaCl to a final concentration of 4 mM and 4% (by volume) of DMSO. This DMSO concentration has been observed in the NMR experiments not destabilizing Gai3 or interfering with IGGi-11 binding.

### 3.2.3 Gβ1γ2 dimer

Human Gβ1γ2 (hereafter Gβγ), was produced in insect cells by our collaborator Prof. Alan Smrka at the University of Michigan. Pure fractions from IMAC containing Gβγ at 1 g/l (22 μM) in 20 mM HEPES pH 8.0, 0.2 mM EDTA, 1 mM DTT, and 300 mM NaCl were shipped frozen. After thawing, protein concentration was determined by absorbance at 280 nm, concentrated to 50 μM, and mixed with a Gai3 solution to a final [Gai3]/[Gβγ]=1.3. An 80 μM Gai3βγ NMR sample (with excess free Gβγ) in NMR buffer with 300 μM GDP was prepared by three cycles of dilution and concentration. Because the affinity of Gα subunits for Gβγ is on the order of nM at room temperature (Sarvazyan et al., 1998), it is assumed that all Gai3 molecules are bound to Gβγ.

### 3.3 Size exclusion chromatography-multiangle light scattering

Size exclusion chromatography-multiangle light scattering (SEC-MALS) is an analytical technique that combines size exclusion chromatography with multiangle laser light scattering detector. SEC is commonly used to separate biomolecules based on size, providing information about the molar mass distribution in a sample by comparing the elution volume with standards (globular proteins of known mass).

However, this procedure is not appropriate for non-globular proteins or those that weakly interact with the matrix of the column (because their elution volumes do not only depend on their molar masses). MALS measures the scattering of light by molecules as they come out of the column. When a sample is illuminated with a beam of light, the particles scatter the light in various directions. The intensity of the scattered light depends on the molar mass and concentration of the scattering particles. The concentration is measured with another detector (absorbance or refraction index, or both). By positioning MALS detectors around the sample, scattered light can be collected from multiple angles and analyzed to determine the biomolecule's molar mass (and hence the oligomeric state), irrespective of its shape or elution volume (Folta-Stogniew & Williams, 1999).

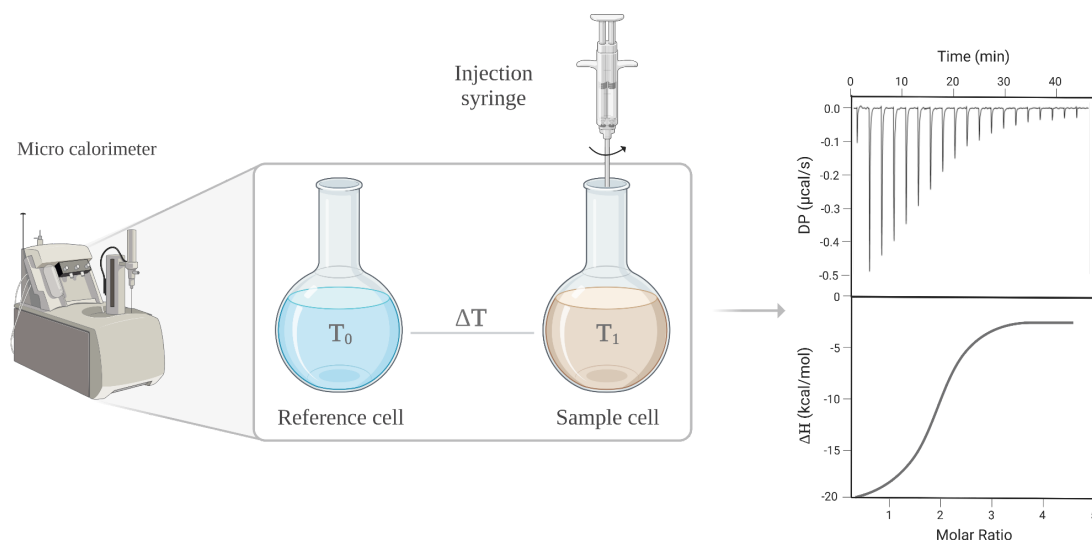
The experiments were conducted at a temperature of 23 °C using a Superdex 200 Increase 10/300 GL column (Cytiva). The column was connected to a DAWN-EOS light scattering detector and an Optilab rEX differential refractive index detector from Wyatt Technology. Prior to the experiments, the column was equilibrated with pH 7.4 PBS (10 mM phosphate, 140 mM chloride, 153 mM sodium, and 4.5 mM potassium) that had undergone 0.1 µm filtration. A sample of 250 µl with sGai3 protein at 1.27 g/l was injected into the system, and the run was performed at a flow rate of 0.5 ml/min. Data acquisition and analysis were conducted using ASTRA software (Wyatt). To calibrate the SEC-MALS systems, a sample of Bovine Serum Albumin (BSA) at a concentration of 2 g/l in the same buffer was used. Through numerous measurements on BSA under similar conditions, we estimate that the experimental error in determining molar mass is approximately 5%.

### **3.4 Isothermal titration calorimetry**

Isothermal titration calorimetry (ITC) is a technique used to study the interactions between biomolecules by measuring the heat exchange caused by the binding reaction. Thermodynamics quantities can be calculated, and thus the affinity and stoichiometry, of the interaction (Srivastava & Yadav, 2019).

In an ITC experiment, two cells are used: the reference cell and the sample cell (**Figure 12**). The reference cell contains a solution in which no binding event occurs and serves as a baseline for the heat measurement. The sample cell contains the biomolecule of interest. The microcalorimeter uses electricity to heat the cells and keep them at the same constant temperature. The ligand, which is the molecule being

titrated, is added in small increments to the sample cell using a stirring syringe. As the ligand binds to the target molecule, heat is released or absorbed, changing the temperature of the sample cell. Feedback from the temperature sensors triggers the heater to adjust its power so that the cell returns to the set temperature (identical to the reference cell). The change in electricity power is related to the heat exchange, which, at constant pressure, is the enthalpy change. The titration allows calculating other thermodynamic quantities, and thus the affinity and stoichiometry of the interaction.



**Figure 12.** Schematic representation of the principle behind ITC. The micro calorimeter monitors the temperature of the sample cell and reference cell, allowing for the measurement of heat changes during the binding process. The upper panel illustrates the differential power as a function of time, displaying a series of peaks that correspond to each injection, indicating the binding events between the macromolecule and ligand. The lower panel presents the plot of enthalpy change as a function of the molar ratio of the macromolecule to the ligand.

Measurements were performed using a MicroCal PEAQ-ITC calorimeter (Malvern) at 25 °C. Prior to the measurements, proteins and peptides were dialyzed against 10 mM HEPES pH 7.0, 10 mM MgCl<sub>2</sub>, 1 mM TCEP. Subsequently, GDP was added to both samples and dissolved in a small volume of the dialysis buffer to a final concentration of 300 μM. IGGi-11 was not dialyzed, the concentrated stock was diluted in the dialysis buffer with 300 μM GDP.

To measure IGGi-11 binding to Gαi3, the sample cell was loaded with 50 μM protein and the syringe contained 1 mM IGGi-11. To measure binding to sGαi3, the sample

cell was loaded with 20  $\mu\text{M}$  protein for GIV<sub>1671-1691-DY</sub> and 41  $\mu\text{M}$  sGai3 for IGGi-11, and the syringe was loaded with either 200  $\mu\text{M}$  GIV<sub>1672-1693-DY</sub> or 500  $\mu\text{M}$  IGGi-11. The experimental setup involved a series of 19 injections, each consisting of 2  $\mu\text{l}$ , with a time spacing of 150 s and a stirring speed of 750 rpm. As a control, dilution heats for the syringe reactant were determined by carrying out identical injections of the reactant into the sample cell loaded with buffer. The electrical power required to maintain a constant temperature in the reaction cell after each injection was recorded over time. The binding isotherms were fitted to a 1:1 binding model, taking into account dilution effects, using the MICROCAL PEAQ-ITC analysis software (Malvern). The number of binding sites was, however, an adjustable parameter. A number close to 1 indicates that the interaction model is appropriate.

### 3.5 Nuclear magnetic resonance spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful technique used to study the structure, dynamics, and interactions of molecules. The technique relies on the fact that some atomic nuclei possess a property called spin, which can only be appropriately described with the formalism of quantum mechanics, but for some aspects of its behavior classical mechanics analogies are useful. Nuclei with spin behave as magnetic moments in the presence of an external magnetic field, and their relative alignment is quantified, with discrete states with different energies. For spin 1/2 nuclei two states are possible: parallel alignment (low energy) and antiparallel alignment (high energy). By applying a radiofrequency pulse matching the difference in energy, transitions and coherences (resonances) can be induced. System relaxation back to equilibrium is governed by mechanisms strongly dependent on molecular motion and occur in slow time scales compared with other spectroscopies (see below). Molecules in solution move fast and the NMR spectra are simplified with respect to the solid state, with much narrower spectral lines. Correlations between coupled spins can be measured to extract structural information: scalar couplings (through covalent bonds) and dipolar couplings (between spins close in space). For NMR spectroscopy on proteins used to study molecular binding events it is very useful to register two-dimensional spectra that correlate the frequencies of <sup>1</sup>H and <sup>15</sup>N nuclei (both spin 1/2) of atoms that are directly bonded, like those in the protein backbone (Mlynárik, 2017).

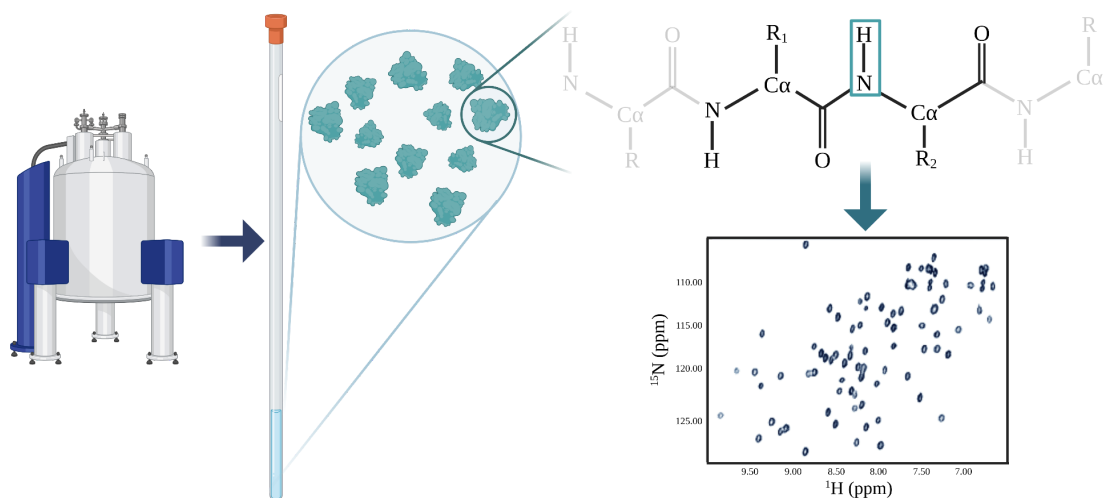


The most frequently used such spectra are recorded as heteronuclear single quantum coherence (HSQC) experiments (**Figure 13**). In this experiment proton frequencies are detected directly and  $^{15}\text{N}$  frequencies indirectly, through their scalar coupling with the covalently bound proton. The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum shows correlations of backbone amide groups, one per residue (except for prolines and the N-terminus), plus a few other correlations from Trp, Asn, Gln, His, Arg, and Lys side chains. Because the hydrogen in all these NH groups can form hydrogen bonds with the water molecules of the buffer, their observation depends on the pH and the temperature of the experiment as well as on the molecular structure (which may sequester them from the solvent (Cavanagh et al., 2010)).

The natural abundance of  $^{14}\text{N}$  is 99.6%, but this isotope has unfavorable NMR properties, leading to broad NMR signals, making  $^1\text{H}$ - $^{14}\text{N}$  HSQC NMR experiments not useful for proteins. The natural abundance of  $^{15}\text{N}$  isotope is very low (0.4%), and only concentrated samples (above 1 mM) may yield useful HSQC experiments (in a few hours). For large proteins sensitivity and resolution is further reduced by the slow molecular tumbling in solution at regularly used temperatures, which broadens the NMR signals. For these reasons, uniform  $^{15}\text{N}$  enrichment is currently done for almost any protein to be studied by NMR. For large proteins (above 20 kDa) it is useful to substitute non-labile protons by deuterons (see below), and it is indispensable for very large proteins. This is useful for proteins up to about 100 kDa; beyond that size other strategies are necessary. Deuteration is not necessary for intrinsically disordered proteins as fast local dynamics narrows the lines. Uniform deuterium enrichment is usually achieved by feeding the bacteria producing the protein with deuterated glucose and growing them in deuterated water (with  $^{15}\text{N}_2\text{H}_4\text{Cl}$  for simultaneous  $^{15}\text{N}$  enrichment). Most of the labile deuterons are exchanged with protons during the purification of the protein in protonated buffers. Therefore, most or all of the nitrogen atoms are bound to protons after protein purification, but some deuterons, like those sequestered from the solvent in the hydrophobic core, may not be sufficiently exchanged for proton NMR signal observation. Selectively  $^{15}\text{N}$  enrichment in certain amino acid types might be useful to simplify the  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra and disambiguate signal assignments. This is achieved by feeding the bacteria with the 20 amino acids, one (or more) of them with  $^{15}\text{N}$  isotope enrichment.

For molecular recognition studies of the 41 kDa G<sub>ai3</sub> protein by NMR it was necessary to prepare uniformly enriched <sup>2</sup>H-<sup>15</sup>N samples and register <sup>1</sup>H-<sup>15</sup>N correlation spectra of the TROSY type (see below) at high magnetic field (with a <sup>1</sup>H frequency of 800 MHz). With a 50 μM sample in a 5 mm NMR tube, it was possible to measure a useful TROSY spectrum in 3 h at 30 °C in Bruker Avance 800 MHz spectrometer and a TCI-cryoprobe with a z-gradient. Spectra were acquired and processed with TopSpin and analyzed with Sparky (Lee et al., 2015).

The residue specific assignment of the GDP-G<sub>ai3</sub> NMR spectrum was previously reported (Mase et al., 2014) at 30 °C using an almost identical buffer. This assignment could be easily transferred to most signals observed in the <sup>1</sup>H-<sup>15</sup>N TROSY spectra. The assignment of the GTP<sub>γ</sub>S-G<sub>ai3</sub> form was reported in the same buffer but at pH 6.5 (Mase et al., 2012), and it could also be transferred for most of the signals. The reported assignment is extensive but not complete: only 85 % of the amide protons were identified, possibly due to unfavorable local dynamics and/or incomplete deuteration exchange to protons. The percentage of amide protons assigned in the G<sub>ai3</sub>-ligand studies in this thesis is smaller, 71 %, probably due to the low protein concentration used, incomplete deuterium exchange, and signal overlap in some regions of the two-dimensional TROSY.



**Figure 13.** Schematic representation of the Heteronuclear Single Quantum Coherence (HSQC) NMR experiment for studying proteins. The HSQC experiment correlates the chemical shifts of <sup>1</sup>H and <sup>15</sup>N nuclei in the protein. By focusing on the amide bonds, which contain an NH bond, unique signals are obtained per residue, creating a distinctive fingerprint for the protein.

### 3.5.1 NMR relaxation and protein backbone dynamics

The coherences that give rise to NMR signals of molecules in solution disappear when the nuclear spin system relaxes back to equilibrium. This occurs in a ms-s time scale. NMR relaxation can be characterized by two time-constants:  $T_1$ , for the nuclear magnetization return to equilibrium, and  $T_2$ , for the nuclear magnetization dephasing. After a time  $T_1$ , the system has progressed about 66 % towards full longitudinal relaxation (return to equilibrium).

Transverse relaxation causes coherences to disappear, and thus signal vanishing. This occurs much faster than longitudinal relaxation for proteins in solution. The  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra are recorded with excitation and acquisition at the proton frequency, the  $^{15}\text{N}$  frequency being indirectly detected. Therefore, the most relevant time constants in practice are those of the protons. The relaxation of the protons is dominated by dipolar coupling to other protons, strongly dependent on the internuclear distance and the external magnetic field (which determines the resonance frequency). In solution, NMR relaxation is strongly dependent on molecular motion, both global (molecular tumbling) and local (internal molecular dynamics), and therefore on molecular size, flexibility, and temperature. For Gai3, under the conditions used in this study, the proton  $T_2$  is estimated to be of a few ms for the most rigid regions. Such a short time makes it impossible to record useful NMR spectra (signals are weak and broad, with insufficient sensitivity and resolution). Proton relaxation can be slowed down by deuteration of the non-labile protons (most carbon-bound protons). By reducing the density of protons in the vicinity of the amide protons,  $T_2$  increases by a factor of 10, making it possible to record spectra with intense and narrow signals at practical protein concentration and NMR time. A specialized  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra named TROSY further improves sensitivity and resolution in large deuterated proteins at high magnetic fields. By exploiting the constructive interference between the two dominant mechanisms in transverse relaxation, TROSY allows for the acquisition of high-quality spectra for large proteins such as Gai3.

The dynamics of the polypeptide chain can be studied by specialized NMR experiments that are sensitive to motions in particular time scales. Local backbone dynamics in the ps-ns time scale can be inspected by measuring the time constants of the  $^{15}\text{N}$  longitudinal relaxation ( $T_1$ , for the nuclear magnetization return to

equilibrium) and transverse relaxation ( $T_2$ , for the nuclear magnetization dephasing), as well as the nuclear Overhauser effect (NOE) on the  $^{15}\text{N}$  upon selective radiofrequency irradiation of  $^1\text{H}$ . For the backbone amide  $^{15}\text{N}$ , these relaxation parameters are dominated by the dipolar coupling to its covalently bound  $^1\text{H}$ , and the  $^{15}\text{N}$  chemical shift anisotropy. Both magnitudes depend on the internuclear distance and the relative orientation of the interatomic bond with respect to the external magnetic field. The internuclear distance can be considered time independent (bond vibrations occur on a fs-ps time scale). The relative orientation varies with time due to global motion (molecular tumbling) and local motion (backbone dynamics). Therefore, the flexibility of the protein backbone at each residue is actually probed. For proteins in solution at regularly used temperatures the more mobile the NH bond the larger the  $T_1$  and the smaller the  $T_2$ . The heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  NOE is always negative (the effect decreases the intensity), but the numerical quantity that is typically used to represent the NOE (ratio of signal intensity in spectra with or without irradiation) is positive and close to 0.9 for rigid backbone sites, and negative for very flexible sites.

$^{15}\text{N}$ -relaxation experiments were conducted on a 125  $\mu\text{M}$  sample of  $^2\text{H}$ - $^{13}\text{C}$ - $^{15}\text{N}$ -Gai3, in 10 mM HEPES pH 7.0, 10 mM  $\text{MgCl}_2$ , 5 mM DTT, 300  $\mu\text{M}$  GDP, 5%  $\text{D}_2\text{O}$ ,  $\text{NaN}_3$  0.01% and 50  $\mu\text{M}$  DSS, on a Bruker AVANCE III 800 spectrometer with a TCI cryoprobe with a z-gradient coil at 30 °C. The heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  NOE was measured, in an interleaved mode, with a relaxation/irradiation time of 10 s. The same conditions were used to record experiments on a 100  $\mu\text{M}$  sample of  $^2\text{H}$ - $^{15}\text{N}$ -Gai3 with 200  $\mu\text{M}$  GIV<sub>C-t</sub> (1:2 molar ratio). Measurements for  $T_1$  and for  $T_2$  time constants were acquired in a fully interleaved mode alternating the corresponding relaxation times between each scan, thus eliminating any long-term instabilities and differential heating effects.  $T_1$  values were derived from the decay in the signal intensities measured in 8 spectra recorded with the corresponding delay in the pulse sequence in the range 0.025-1.6 s (with a total duration of 23 h).  $T_2$  values were derived from the decay in the signal intensities measured in 6 spectra recorded with the corresponding delay in the pulse sequence in the range 0-0.08 s (with a total duration of 19 h). The decay of the peak heights at each time were fitted to a one exponential equation with GraphPad Prism, yielding the rate constants  $R_1$  and  $R_2$  (the inverse of the time constants) and the fitting errors for each residue that could be

reliably measured (those with signal intensity above the noise level and without severe signal overlap).

Relaxation data of GDP- G $\alpha$ i3 were fitted to the Lipari and Szabo model (Lipari & Szabo, 1982) using FAST-Modelfree (Cole & Loria, 2003), which interfaces with MODELFREE version 4.2 (Palmer III et al., 1991). The  $S^2$ ,  $\tau_e$ , and  $R_{ex}$  models of internal motion were evaluated for each amide  $^1\text{H}$ - $^{15}\text{N}$  pair, where  $S^2$  is the generalized order parameter,  $\tau_e$  is the effective internal correlation time, and  $R_{ex}$  is the exchange contribution to transverse relaxation. For this analysis, the crystal structure of sG $\alpha$ i3 was used to calculate the principal components of the inertia tensor of the protein, using the program Pdbinertia (A.G. Palmer III, Columbia University). The overall correlation time was estimated from the ratio of the mean values of  $T_1$  and  $T_2$ . These mean values were calculated from a subset of residues with little internal motion and no significant exchange broadening. This subset excluded residues with NOEs less than 0.65 and also residues with  $T_2$  values less than the average minus one standard deviation, unless their corresponding  $T_1$  values were larger than the average plus one standard deviation (Pawley et al., 2001). The diffusion tensor, which describes rotational diffusion anisotropy, was determined by two approaches (Tjandra et al., 1995; Brüschweiler et al., 1995) with the r2r1\_diffusion program and the quadric\_diffusion program (A.G. Palmer III, Columbia University). The  $^{15}\text{N}$  relaxation was analyzed assuming dipolar coupling with the directly attached proton (with a bond length of 1.02 Å), and a contribution from the  $^{15}\text{N}$  chemical shift anisotropy evaluated as -160 ppm.

### 3.5.2 Protein-ligand binding studies

Chemical shifts ( $\delta$ ) are highly sensitive to the local environment of each residue, making them valuable for studying protein-ligand interactions at an atomic level, even when the affinity is low. The changes in chemical shifts observed between the free and bound states of a protein are commonly referred to as chemical shift perturbations (CSPs) (Williamson, 2013).

In this work, we have employed the following equation to quantify CSPs measured in  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra:

$$CSP = \sqrt{\frac{1}{2} (\Delta\delta_H^2 + \frac{\Delta\delta_N^2}{25})}$$

By analyzing CSPs, we can detect that there is binding, identify the binding site(s) and, in favorable cases, gain insights into the structural changes induced by the ligand on the protein.

CSPs measured along a titration (where the ligand concentration is varied) can be utilized to calculate the affinity of the protein-ligand interaction. In favorable cases, the chemical shift is the weighted average of the values corresponding to the nuclei in molecules in their free-state and in the bound state. Therefore, the CSP is directly proportional to the relative concentration of the bound state. The CSPs measured at each titration point can be fitted to a 1:1 binding model according to this equation:

$$CSP = \frac{K_D + [P] + x[P] - \sqrt{(K_D + [P] + x[P])^2 - (4x[P]^2)}}{2 [P] CSP_{max}}$$

where [P] is the total protein concentration (assumed to be constant), x is the molar concentration ratio ([ligand]/[protein]), K<sub>D</sub> is the dissociation constant, and CSP<sub>max</sub> is the CSP at saturation. The two adjustable parameters are K<sub>D</sub> and CSP<sub>max</sub>.

The experimental condition in which this approach is valid is when the exchange rate between the free and bound states is larger than the difference in the frequencies of the nuclei in both states. For protein <sup>1</sup>H-<sup>15</sup>N signals this is commonly the case when K<sub>D</sub> is larger than 1 μM (assuming an association rate constant limited by molecular diffusion, about 10<sup>-7</sup> s<sup>-1</sup>M<sup>-1</sup>).

Two statistical criteria were used for considering significant CSPs. The most common approach is to define a threshold as the average CSP plus one (or two) standard deviations. This was utilized when selecting residues for calculating the dissociation constant of the Gai3-IGGi-11 interaction or assessing docking results. The second approach defines CSP significance ranges based on the value of the median. This approach was previously used for Gai3 binding study to GIV<sub>C-t</sub> and it was found to be useful to describe the effect on the different regions of the protein. For consistency with that report, this approach is used here to describe the binding of the different ligands. Some protein signals are perturbed upon ligand binding to such extent that their assignment in the bound form cannot be done. In the spectra the signal seems to vanish, because the intensity at the frequencies corresponding to the free state is below the noise, and there is too much ambiguity to transfer the assignment to other signals by the nearest-neighbor approach. The residues that are perturbed in such a

way are classified as having a strong drop in intensity ( $I_r < 0.01$  in the legends of the figures).

NMR spectra on Gai3 and its complexes were registered on a Bruker AVANCE III 800 spectrometer equipped with a TCI cryoprobe and z-gradient coil at 30 °C. The samples, about 400  $\mu$ l in a 5 mm Shigemi tube without plunger, contained Gai3 at different concentrations in 10 mM HEPES pH 7.0, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 5% D<sub>2</sub>O, 0.01% NaN<sub>3</sub>, and 50  $\mu$ M DSS. The solution also included 300  $\mu$ M GDP for all ligands except KB-1753, for which 300  $\mu$ M GTP $\gamma$ S was used instead. GTP $\gamma$ S is a non-hydrolyzable analog of GTP, with one of the three oxygens at the terminal phosphate substituted by a sulfur atom. IGGi-11 binding was measured to both GDP-Gai3 and GTP $\gamma$ S-Gai3. The complexes with G $\beta\gamma$  or with GIV<sub>C-t</sub> were prepared by mixing <sup>2</sup>H-<sup>15</sup>N-Gai3 with G $\beta\gamma$  or with GIV<sub>C-t</sub> in their corresponding purification buffers followed by three cycles of dilution and concentration in NMR buffer with GDP (and then supplemented with <sup>2</sup>H<sub>2</sub>O, NaN<sub>3</sub> and DSS).

For titrations, a few  $\mu$ l of concentrated ligand stock solutions were pipetted into 400  $\mu$ l of Gai3 samples in a 5 mm shigemi tube without plunger, which was capped and mixed by inverting the tube about 10 times after each addition. At least 20 min passed between ligand addition and the start of the TROSY spectrum measurement, and it was assumed that the binding reaction had reached the equilibrium. The duration of each of the TROSY experiments was identical along a titration, but differed between samples. TROSY spectra were recorded for 2-4 h depending on the protein concentration and spectrometer time availability. A titration experiment was completed in less than 2 days. At the end of the titration the protein concentration was diluted by no more than 10%. In some instances, protein precipitation occurred inside the NMR tube upon ligand addition, but in little amount, as judged visually and by the NMR signal intensity in the TROSY spectra. When the ligand stock solution contained DMSO, the concentration of DMSO at the last point of the titration was below 2.5 %. To assess the ability of IGGi-11 to displace the protein or peptide ligands from Gai3, IGGi-11 was added at the end of the titrations to a final concentration between 8 and 32-fold molar excess over the ligand concentration, and DMSO concentration increased up to 3.9 %. In some instances, partial protein precipitation occurred, but a causal relationship between precipitation and DMSO content could not be established. About 20 min after IGGi-11 addition, TROSY spectra were acquired to monitor the changes caused by excess IGGi-11.

TROSY experiments on <sup>15</sup>N Trp Gai3 W258F in both its free form and when bound to Gβγ, R6A, RGS12, GIV<sub>1671-1705</sub>, and IGGi-11, were conducted under the same conditions. However, as the protein was not deuterated, long experiments (about 18 h) were necessary to achieve a useful signal to noise ratio.

**Table 3.** Gai3 and ligand concentrations in the NMR samples used in titrations.

Ligand	Initial [Gai3]	[Gai3]:[ligand]
R6A	38 μM	1:0, 1:0.3, 1:0.6, 1:1.1, 1:3
R6A-1	38 μM	1:0, 1:0.2, 1:0.4, 1:0.8, 1:2, 1:4
GIV <sub>1671-1705</sub>	32.5 μM	1:0, 1:4
RGS12	78 μM	1:0, 1:0.1, 1:0.4, 1:0.8, 1:1.6, 1:4
Gβγ	80 μM	1:0, 1:1.3
KB-1753	13 μM	1:0, 1:0.35, 1:0.67, 1:1.34, 1:2.67, 1:5.38
IGGi-11	65 μM (GDP)	1:0, 1:0.2, 1:0.4, 1:0.8, 1:1.6, 1:3.2, 1:6.4
	70 μM (GTPγS)	1:0, 1:1, 1:4, 1:8, 1:16, 1:35

### 3.5.3 Measurement of guanine nucleotide exchange rates in Gai3

The rate of GDP exchange for GTP in Gai3 was assessed using an NMR-based method (Gebregiworgis et al., 2018) that relies on monitoring changes in NMR signal intensity between GDP-bound and GTPγS-bound states. The analysis focused on the indole signals of W211 and W258 whose chemical shifts are different in both states in a region with little signal overlap, so that intensity quantification is most reliable.

The assay was performed by adding 5 μl of a 70 mM GTPγS stock to a 89 μM GDP-bound <sup>2</sup>H-<sup>15</sup>N-Gai3 sample in 400 μl of 10 mM HEPES pH 7.0, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 5% D<sub>2</sub>O, 0.01% NaN<sub>3</sub>, and 50 μM DSS, in a 5 mm Shigemi tube (10-fold molar excess of GTPγS). A total of 30 NMR spectra, each with a duration of 28 min 54 s, were sequentially acquired at 20°C after the addition of a few μl of a concentrated stock of GTPγS in the same buffer and mixing by tube inversion. 11 minutes passed between addition and the start of the first TROSY. For data analysis each measurement was assigned to the time at the midpoint of the corresponding TROSY plus 11 minutes.

The decay or growth in the indole signals was quantified by measuring the peak height using Sparky, and fitted to equations for one exponential decay or for one phase association using GraphPad Prism 9.



### 3.6 Crystallography

X-ray crystallography is a powerful method for determining the precise three-dimensional structure of proteins at the atomic level. The process begins by purifying and concentrating the protein of interest to obtain a high-quality crystal.

During protein crystallization, purified proteins are mixed with a suitable precipitant or a combination of precipitants under controlled conditions of pH, temperature, and concentration. The aim is to find the optimal conditions for the formation of well-ordered crystals by mixing purified proteins with suitable precipitants.

Various methods and techniques have been developed to aid in protein crystallization, including vapor diffusion, batch crystallization, microbatch, and high-throughput screening approaches. In this project, the sitting drop method was employed. It involves combining the protein solution with a precipitant and placing the resulting drop directly onto a well of a crystallization plate adjacent to a reservoir containing a larger volume of the precipitant solution. To prevent evaporation, the drop is covered with a lid or sealing tape. Over time, the drop and the reservoir reach equilibrium through the diffusion of water vapor, allowing for the establishment of a stable environment within the drop that can potentially lead to protein crystallization.

Protein crystallization involves nucleation and growth. Nucleation is the formation of tiny crystals, while growth refers to their increase in size over time. Factors such as protein concentration, precipitant conditions, and temperature influence both nucleation and growth. Successful protein crystallization requires obtaining well-formed nuclei and promoting their growth into larger, high-quality crystals. For that purpose, we have employed two strategies: the addition of additives and seeding. The addition of specific additives, such as salts, organic solvents, or small molecules, has been used to improve protein solubility, stability, and crystal quality. We specifically utilized detergents as additives. Detergents, being amphipathic molecules, have the ability to solubilize membrane proteins or proteins with hydrophobic regions. They contribute to stabilizing the protein structure, preventing aggregation, and promoting the formation of well-ordered crystals.

Seeding involves introducing a seed crystal, which can be a small crystal or cluster of crystals, into the protein solution. This seed crystal acts as a template, providing a surface for the protein molecules to align and form nuclei. It serves as a starting point

for the growth of new crystals, facilitating the nucleation process. By overcoming the barriers to nucleation, seeding increases the likelihood of successful crystal growth and the production of well-formed crystals.

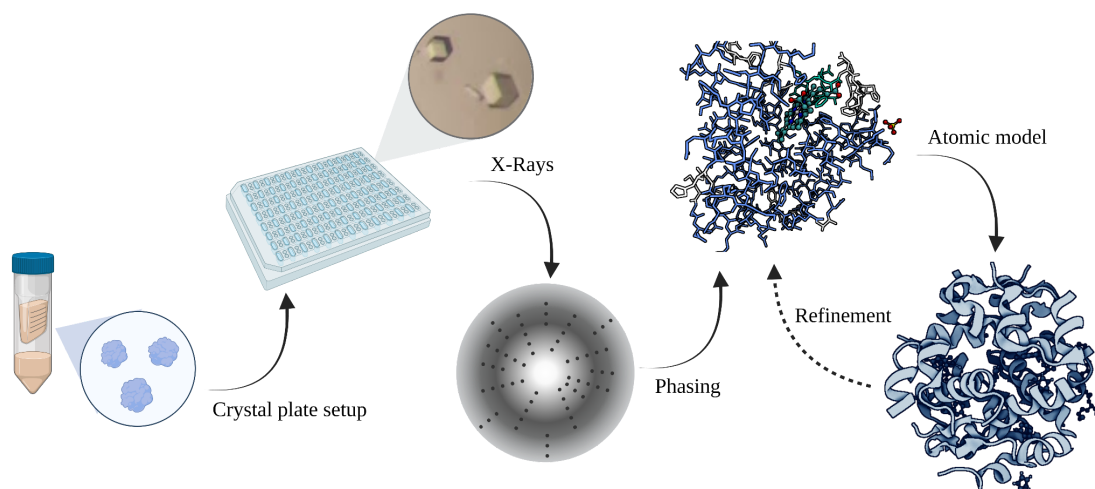
Once the crystal is formed, it is then carefully mounted and subjected to a focused X-ray beam. As the X-rays interact with the electrons present in the crystal lattice, they undergo diffraction, resulting in a distinctive pattern of scattered X-rays.

To extract meaningful information from the diffraction pattern, the crystal must be carefully aligned and rotated in different orientations to collect data from various crystallographic planes. Multiple diffraction images are recorded, covering a wide range of angles, which are used to determine the intensities and positions of the diffracted X-rays. This data is then subjected to analysis using mathematical algorithms and computational techniques. Through these calculations, an electron density map is generated, which represents the distribution of electrons within the crystal lattice. The electron density map is then interpreted to determine the positions of individual atoms in the protein. Through multiple rounds of refinement, the three-dimensional structure of the protein is deciphered, revealing the precise arrangement of atoms (**Figure 14**).

Crystallization trials were performed for sGai3, in its free form and when bound to GIV<sub>1672-1693-DY</sub> or IGGi-11 at a molar ratio of 1:2. Purified sGai3 was in 50 mM Tris pH 8.0, 300 mM NaCl, 5% Glycerol, 20 μM GDP and 1 mM DTT. To decrease the NaCl concentration for crystallization trials, the sample was diluted 2-fold with distilled water and concentrated to 10 g/l.

Initial crystallization trials were performed at 22 °C using the sitting-drop vapor-diffusion method with commercial screening solutions, including JBScreen Classic and Wizard Classics I–IV (Jena Bioscience, Jena, Germany), in 96-well sitting-drop plates (Swissci MRC; Molecular Dimensions, Suffolk, England). Drops were set up by mixing equal volumes (0.2 μl) of protein-containing solution and reservoir solution using a Cartesian Honeybee System (Genomic Solutions, Irvine, USA) nanodispenser robot and equilibrated against 50 μl reservoir solution. Crystals of sGai3 in the presence of IGGi-11 were successfully obtained within one day using a crystallization condition consisting of 0.1 M sodium MES at pH 6.5, 20% PEG 8000, and 0.2 M ammonium sulfate. Similarly, crystals of sGai3 in the presence of GIV<sub>1672-1693-DY</sub> were obtained within one day using a condition containing 18% PEG

8000, 0.5 M lithium sulfate, and EDTA. However, crystals did not grow with sGai3 alone. For data collection, the crystals were harvested in a precipitant solution containing 30% (v/v) glycerol or ethylene glycol, and were flash-cooled in liquid nitrogen. While single well-diffracting crystals were obtained for sGai3 grown with IGGi-11, the resolution for the crystal grown with GIV<sub>1672-1693-DY</sub> was too low and the data were not processed.



**Figure 14.** X-ray crystallography process for protein structure determination. The process involves several steps: crystallization, X-ray exposure, diffraction pattern analysis, extraction of atomic positions and construction of a three-dimensional model.

X-ray diffraction data were collected at the ALBA Synchrotron (Cerdanyola del Vallès) BL13 XALOC beamline. Data indexing and integration were performed using AutoProc and Staraniso from Globalphasing, along with XDS for scaling and merging (Vonrhein et al., 2011; Tickle et al., 2018; Kabsch, 2010). The CCP4 suite programs POINTLESS and AIMLESS were utilized for further processing and analysis (Evans, 2006; Evans & Murshudov, 2013; Winn et al., 2011). The structure was solved by molecular replacement using the structure of the Gai3 from *Rattus norvegicus* bound to GIV peptide (PDB ID: 6MHF, Kalogiropoulos et al., 2019) with Phaser (McCoy et al., 2007). The initial model was first refined using Phenix-refine (Adams et al., 2010) and alternating manual building with Coot (Emsley et al., 2010). The final model was obtained by repetitive cycles of refinement. The model's quality and validation were assessed using MolProbity (Chen et al., 2010), and protein structure figures were generated using PyMOL (Schrödinger & DeLano, 2020).

Interestingly, the diffraction data obtained from the crystals of sGai3 in the presence of IGGi-11 resulted in the unexpected resolution of the structure of free sGai3. To further explore the crystallization of the protein complexes, the sGai3 crystals obtained previously were used as seeds for new crystallization assays. The assays involved the use of an Oryx4 (Douglas Instruments, East Garston, UK) crystallization robot and a protein concentration of 10 g/l. Crystals with IGGi-11 were successfully grown in a crystallization condition consisting of 0.1 M MES pH 6.5, 30% PEG 5000 MME, and 0.2 M ammonium sulfate. Crystals with GIV<sub>1672-1693-DY</sub> were obtained in a condition containing 0.1 M MES pH 6.5, 20% PEG 4000, and 0.6 M NaCl. Free sGai3 did not form crystals. For data collection, the harvested crystals were immersed in a precipitant solution containing 13% and 30% glycerol for the sGai3-IGGi-11 and sGai3-GIV<sub>1672-1693-DY</sub> crystals, respectively, before being flash-cooled in liquid nitrogen. Once again, the diffraction data obtained from the crystals grown in presence of IGGi-11 resulted in the resolution of the structure of free sGai3. However, the resolution of this data was found to be lower compared to the previously obtained structure. As a result, the decision was made not to refine the structure derived from the new data and to retain the previous structure instead.

### **3.7 *In silico* docking of IGGi-11 on sGai3**

Computational modeling can provide valuable insights into the structure of proteins and their complexes when the experimental methods fail, or when they are too difficult or too costly to apply. When no high-resolution structure of the protein is available, a model can be built based on multiple sequence alignment and knowledge gathered from the database of protein structures, as AlphaFold does (Jumper et al., 2021). This approach is still not yet useful to reliably model protein-protein complexes (although there are favorable cases), and not yet implemented for protein-nucleic acid or for protein-small compound complexes. There are several methods to dock a small molecule on a protein structure implemented in computational tools that run on web servers or as stand-alone programs. They can be all considered as an interplay between a search algorithm and a scoring function. Those used in this work are briefly described below.

HADDOCK (High Ambiguity Driven Biomolecular DOCKing) incorporates experimental data as ambiguous interaction restraints (AIRs) to guide the docking. These data may come from NMR measurements or other experiments (such as

comparative mutant protein analysis), and are implemented as ambiguous distance restraints between protein amino acids and ligand amino acids or atoms. The backbone of the protein is kept rigid but side chain rotamers at the interface are explored. The small organic molecule is represented by its corresponding rotamer library. The scoring is based on intermolecular energy terms, which helps identify the most likely complex configurations (Dominguez et al., 2003).

SiteMap identifies potential ligand binding sites on a protein structure. It is fast, but it does not dock a particular ligand (Halgren, 2009). It can be used as a classifier of potential binding sites to selectively dock ligands by other more time-consuming methods.

AutoDock4 does automated docking using a semiempirical free energy force field and selective protein flexibility, and provides relative binding energy predictions (Morris et al., 2009).

AutoDock Vina (Trott & Olson, 2010) is similar to AutoDock4 but much faster, and tends to provide more accurate complexes but less precise binding energies (Nguyen et al., 2020). For virtual screening of ligand libraries, the computational efficiency of Vina makes a large difference. For docking a single molecule, the two methods may yield comparable results, depending on the polarity of the binding site (Vieira & Sousa, 2019).

### **3.7.1 Docking with HADDOCK**

HADDOCK was used to model the sGai3-IGGi-11 complex guided by NMR information (CSPs) as AIRs. The crystal structure of sGai3 was used as the receptor, while a library of rotamers generated by HADDOCK from the IGGi-11 chemical structure (in smile format) was used as the ligand.

The ambiguous interaction restraints (AIRs) were derived from the NMR measurements. Those residues that experienced significant backbone amide chemical shift perturbations were considered to be close to any of the atoms of the ligand. These residues are called “active” by HADDOCK.

### **3.7.2 Docking with AutoDock**

Because the scoring in this docking is based on computing a binding free energy using a free energy force field, the crystal structure was adapted for this force field.

The three N-terminal residues (GAK) and three C-terminal residues (LKE) that were not visible in the electron density of sGai3 were modeled and capped with an acetyl group at the N-terminus and an N-methyl amide group at the C-terminus using the PyMOL Builder tool (Schrödinger & DeLano, 2020). Hydrogen coordinates for the protein and the GDP were generated utilizing AmberTools20 (Case et al., 2020), and a Mg<sup>2+</sup> ion was added based on a structural alignment with the crystal structure of bovine Gai3 bound to GDP (PDB 1TAG), which contains a Mg atom.

The LigPrep wizard in Maestro (Schrödinger, LLC, 2020) was used to create a 3D structure of IGGi-11 from the SMILES file. The protonation state was determined in water at pH 7 ± 2 and the energy minimized, resulting in a net charge of -2 due to the full ionization of the two carboxylic acid groups. The molecular geometry was further refined at the density functional theory calculations and the atomic partial charges were also calculated.

An initial search of pockets in sGai3 was done with SiteMap, as implemented in the molecular visualization environment Maestro (Schrödinger), identifying two possible binding sites. Docking was then conducted using AutoDock Vina 1.1.2 and AutoDock 4.2 with the 3D grid positioned at the  $\alpha$ -carbon of Val225 to include the two possible binding sites (but not including the entire surface of the protein). PDBQT files were generated from the previously described structures, the protonation states at pH 7 ± 2 were calculated, the atomic charges were assigned, and non-polar hydrogens were merged with their covalently bound non-hydrogen atoms. In AutoDock Vina, 20 solutions were generated, and in AutoDock 4, 100 posing runs were performed to ensure adequate conformational sampling.

Molecular Dynamics (MD) simulations were performed to evaluate the stability of Gai3-IGGi-11 complexes. A simulation box was constructed to ensure a minimum distance of 10 Å between any protein atom and the box edges. The system was solvated with TIP3P water molecules and Na<sup>+</sup> ions were added to neutralize the charges. Simulations of 1.0 or 1.5  $\mu$ s were carried out at 300 K in the NPT regime using the Amber20 MD engine (Case et al., 2020). A 1.0  $\mu$ s simulation of the protein without IGGi-11 was also performed for comparison.

Trajectories from the MD simulations were analyzed using Amber20 and VMD 1.9.3. (Humphrey et al., 1996). Python scripts were utilized to compute and plot the Root Mean Square Deviations (RMSD) and Root Mean Square Fluctuations (RMSF). The

assignment of secondary structure at each frame of the simulation was generated using the DSSP (Definition of Secondary Structure of Proteins) method (Kabsch & Sander, 1983). This information was then used to generate a 3D plot, where the residue number is plotted against time, with each residue assigned a specific secondary structure state, which can provide insights into folding/unfolding dynamics and conformational changes.

### **3.8 Cell signaling stimulation, cell lysis, and immunoblotting**

MDA-MB-231 cells come from an epithelial human cell line established from a triple negative metastatic mammary adenocarcinoma. It is one of the most commonly used breast cancer cell lines in research and has very high expression levels of GIV (between 20 and 50-fold GIV mRNA) compared with non-metastatic breast cancer cells or normal epithelial cells (Dunkel et al., 2012). These cells were used in the screening for GIV inhibitors, IGGi-11 reducing their migration and not viability (Zhao et al., 2023).

In the cellular assays, a modified IGGi-11 compound was used. When the carboxylate groups of IGGi-11 are esterified with methyl groups the modified compound, known as IGGi-11me, is more cell permeable and inside cells undergoes esterase-mediated hydrolysis (by endogenous esterases) becoming IGGi-11. Treatment with IGGi-11me attenuated both cell proliferation in three-dimensional cultures of MDA-MB-231 cells and tumor growth in mice injected with tumor cells in the flank (Zhao et al., 2023).

Four-hundred thousand MDA-MB-231 cells or GIV-depleted MDA-MB-231 cells (by means of a small hairpin RNA that blocks GIV expression) were seeded per well on a 60 nm dish. After 24 hours, the media was replaced with DMEM containing 0.5% FBS and 100  $\mu$ M IGGi-11me in DMSO, or DMSO as a control, and the cells were starved for approximately 16 hours in 1.5 ml of DMEM. Following starvation, the cells were stimulated with 1.6 nM EGF for 0 minutes or 5 minutes. The stimulation reactions were terminated by washing the cells three times with ice-cold PBS and adding 100  $\mu$ l of RPPA lysis buffer (Tissue Protein Extraction Reagent (TPER) Buffer, 0.5 M NaCl) supplemented with protease inhibitor cocktail and phosphatase inhibitors before harvesting by scraping.

The whole cell lysates were cleared by centrifugation, quantified using the Bradford method, and then boiled in Laemmli sample buffer for 5 minutes. The proteins were

separated by SDS-PAGE and transferred to PVDF membranes for 2 hours. The PVDF membranes were blocked with TBS supplemented with 5% non-fat dry milk, and then sequentially incubated with primary antibodies. In order to visualize both the total protein bands and the phosphorylated protein bands, different species of antibodies were used. For the detection of total proteins (GIV, Akt, PRAS40, and Tubulin), rabbit antibodies were employed, while for the detection of phosphorylated proteins (pAkt and pPRAS40), mouse antibodies were used. Secondary antibodies utilized were goat anti-rabbit conjugated to AlexaFluor 680 for total protein detection and goat anti-mouse conjugated to IRDye 800 for phosphorylated protein detection.

Infrared imaging of immunoblots was performed using an Odyssey CLx infrared imaging system. Akt and PRAS40 activation were determined by calculating the phospho-protein/total-protein ratio and normalizing it to the maximum activation in each experiment. The images were processed using ImageJ software (NIH) and assembled for presentation using Photoshop and Illustrator software (Adobe).



# Results





---

## 4.1 Characterize the binding of IGGi-11 to Gai3

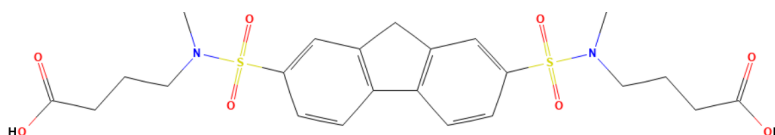
### 4.1.1 GIV fragments containing the GBA motif bind similarly to Gai3

Previous studies showed that the 211-residue long C-terminal region of GIV binds to the G-domain of GDP-Gai3. Perturbations in the NMR signals of many residues in the G domain of Gai3 (but not in the helical domain) indicated closeness to GIV, conformational change induced by GIV binding, or both. Residues within the  $\alpha 3$ , switch II region, and  $\alpha 3$ - $\beta 5$  loop of Gai3 exhibited pronounced perturbations in their NMR signals, suggesting a direct involvement in the interaction. Mutagenesis studies supported that the cavity formed by these regions is the binding site for GIV. Residues involved in GDP binding also displayed substantial chemical shift perturbations, suggesting that the Gai3-GIV interaction destabilizes the binding of phosphate. Overall, these results suggested a mechanism for facilitating nucleotide exchange and activation of Gai3 by GIV (de Opakua et al., 2017).

A detailed analysis of the interaction of GDP-Gai3 with the GIV<sub>1671-1705</sub> fragment has been done here by NMR. This peptide binds Gai3 with a  $K_D$  of  $0.56 \pm 0.01 \mu\text{M}$  at room temperature in fluorescence polarization assays (DiGiacomo et al., 2017). Peptide binding caused perturbations in the Gai3 NMR signals that were very similar to those caused by the C-terminal region of GIV (**Figure 16**). Similar results had been previously reported for the GIV<sub>1671-1696</sub> fragment, confirming that peptides between 19 and 35 residues containing the GBA motif, recapitulate the Gai3 binding properties of GIV (de Opakua et al., 2017).

### 4.1.2 IGGi-11 binds to the same region of GDP-Gai3 that interacts with GIV

The druggability of the Gai3-GIV protein-protein interaction was demonstrated by the dose-dependent inhibition with the small molecule NF023 (DiGiacomo et al., 2017). But NF023 is not cell permeable and it also binds other non-related proteins. A new screening of a library of 200,000 small organic molecules, considering affinity and specificity for Gai3, commercial availability, and effect on cell migration and viability, identified IGGi-11 (Inhibitor of GIV-Gai3 interaction number 11) as a compound with the desired properties (Zhao et al., 2023). IGGi-11 is symmetrical, consisting of a 9H-fluorene core, substituted at positions 2 and 7 by N-(3-carboxypropyl)-N-methylsulfamoyl groups (**Figure 15**). The core is highly hydrophobic but the substituents are polar, and expected to have a negative charge at their ends at neutral pH (the pKa of butyric acid is 4.8 at 25 °C).



**Figure 15.** Molecular Structure of IGGi-11, with nitrogen atoms highlighted in blue, sulfur atoms in yellow, and oxygen atoms in red.

To study the interaction between Gai3 and IGGi-11, an NMR titration experiment was performed at 30 °C. Small volumes of a concentrated stock of IGGi-11 in DMSO were stepwise added to a <sup>2</sup>H-<sup>15</sup>N-labeled GDP-Gai3 sample in aqueous buffer and recording a 3-hour long TROSY spectrum after each addition.

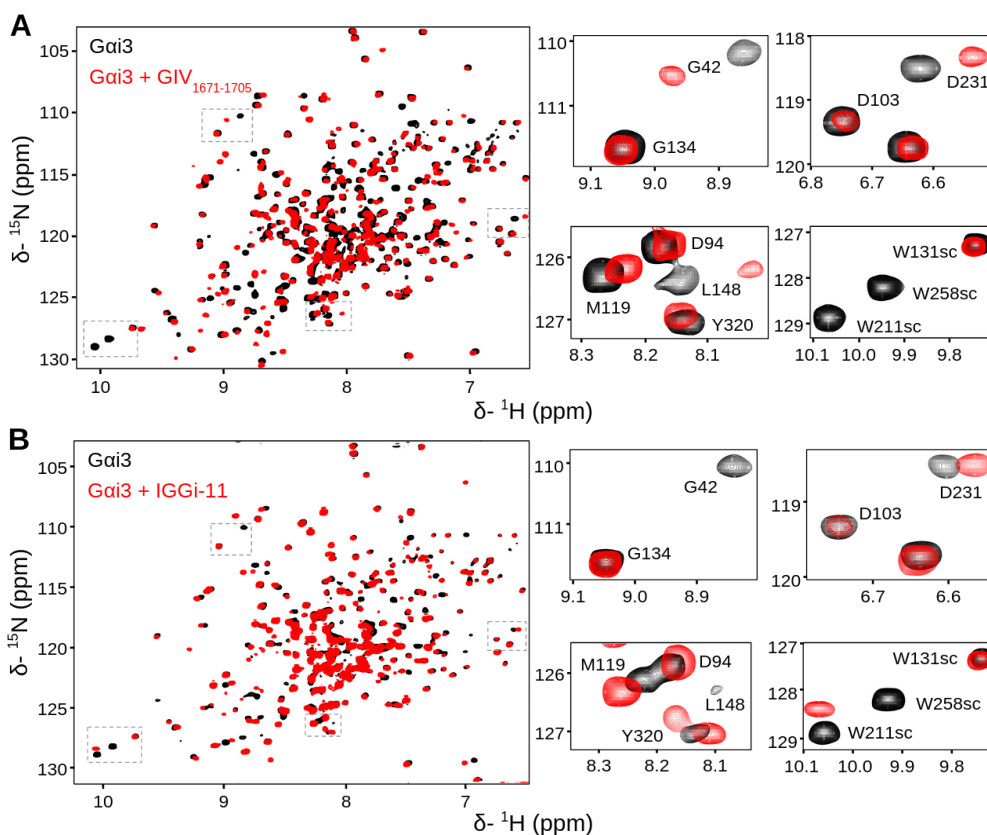
The chemical nature of the GIV peptide and IGGi-11 is very different, and they are expected to cause different signal perturbations when bound to Gai3, but if they both bind to the same (and single) site, the set of perturbed residues should be similar (**Figure 16**). For instance, residues associated with nucleotide binding (G42) or located in the switch regions (W211sc, D231) or in the α3-β5 loop (W258sc) are perturbed by both ligands.

The plot of the CSPs caused by IGGi-11 along the Gai3 residue number exhibits a remarkable resemblance to the plot corresponding to GIV<sub>1671-1705</sub> (**Figure 17**). This is further supported by mapping these perturbations onto a structural model of Gai3.

The helical domain is largely unperturbed in both complexes, indicating that the two ligands interact with the G-domain only, and they bind approximately to the same site. Previous binding assays, involving pull-down experiments of Gai3 with GST-GIV<sub>1671-1755</sub> fusion and fluorescence polarization measurements using fluorescently labeled GIV<sub>1671-1701</sub>, showed that IGGi-11 prevented the interaction. These findings and the NMR data strongly suggest that IGGi-11 acts as a competitive inhibitor of the Gai3-GIV<sub>1671-1705</sub> interaction. But there are many Gai3 residues in the G-domain showing significant CSPs, probably because of conformational changes occurring upon binding, not only due to ligand proximity.

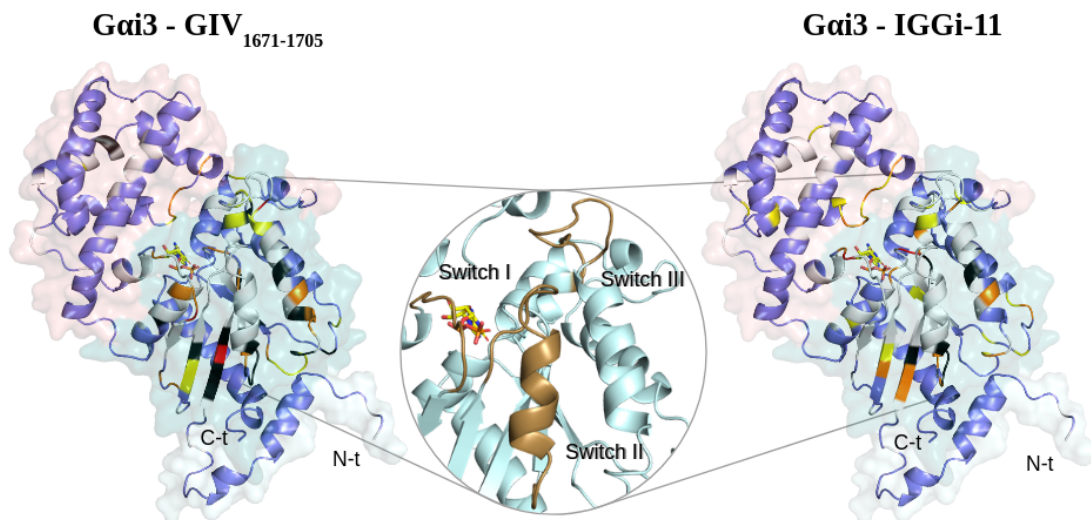
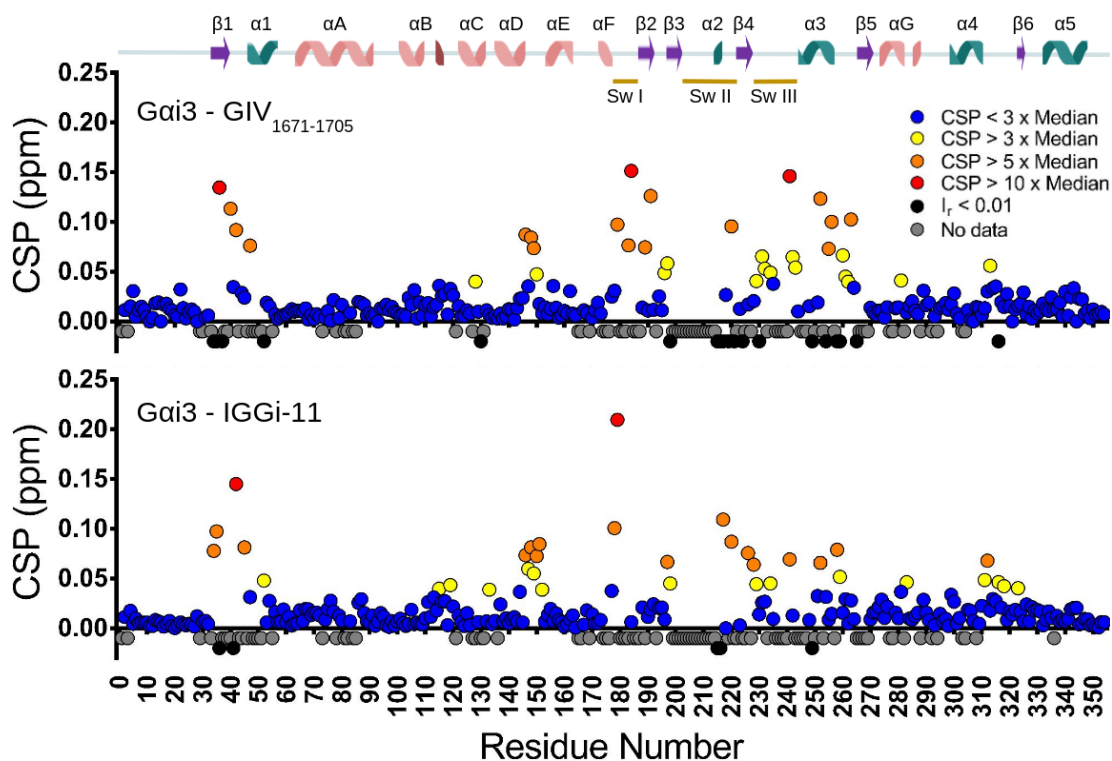
The GIV binding site is known to be the cavity formed by helix α3, the switch II region, and the α3-β5 loop (see below). However, determining with precision that IGGi-11 binds to the same cavity as GIV based solely on NMR data is not straightforward due to the lack of information for several residues around this cavity. Some NMR signals from these residues are either absent or unassigned, particularly in the switch II region. This may be due to local dynamics unfavorable for NMR

signal observation, incomplete deuterium-proton exchange during protein sample preparation, or strong signal overlap.



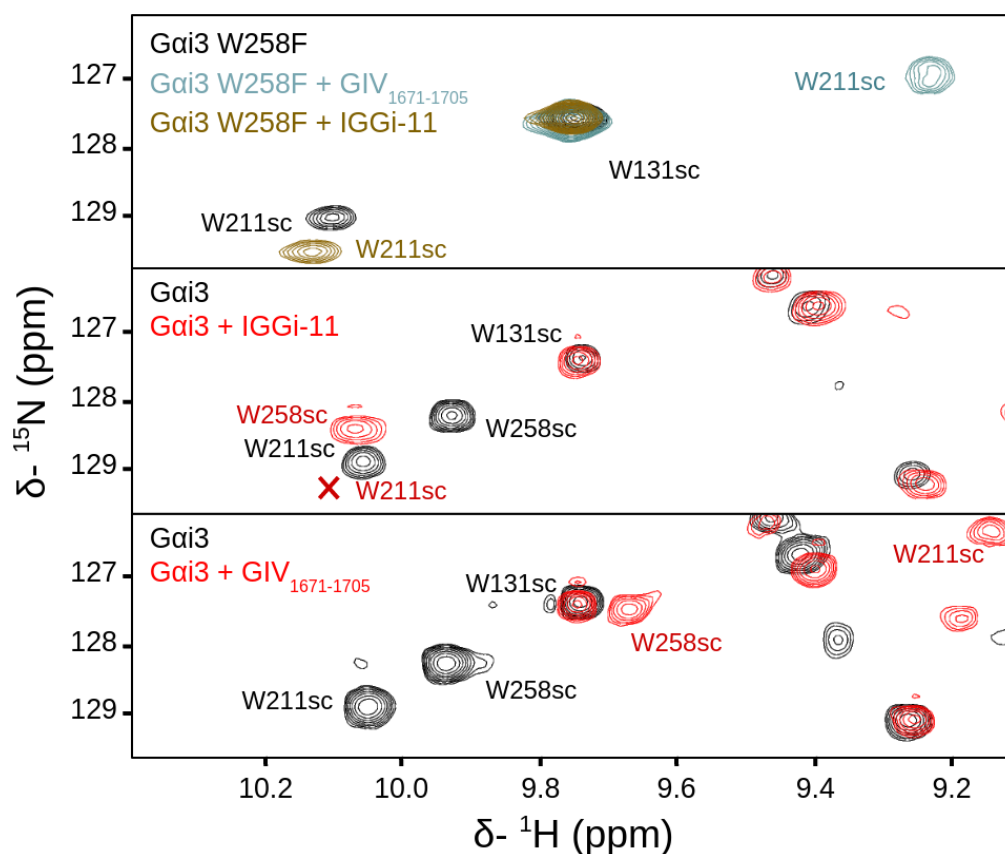
**Figure 16.**  $^1\text{H}$ - $^{15}\text{N}$  TROSY NMR spectra of GDP-*Gai3* in the presence of two different ligands. (A) Overlay of the spectra in the absence (black) and presence (red) of GIV<sub>1671-1705</sub> at a 1:4, molar ratio. (B) Overlay of the spectra of *Gai3* in the absence (black) and presence (red) of IGGi-11 at a 1:6.4, molar ratio. The panels on the right-hand side are zoomed-in selected regions.

The presence of three tryptophan residues in *Gai3*, with side chain NMR signals observed in a region of the spectrum with few other signals, make them very useful to rapidly evaluate ligand binding, albeit with a small spatial resolution. These tryptophan residues are W131 (helical domain), W211 (switch II region); and W258 ( $\alpha$ 3- $\beta$ 5 loop). The assignment of the NMR signal of the side chain of W131 in GDP-*Gai3* could be easily transferred to the complexes with GIV or IGGi-11 (because it was almost unperturbed), but the assignment of W211 and W258 signals was ambiguous, and could not always be solved by tracking their signals along a titration because slow exchange caused some signals to decrease in intensity rather than shifting.



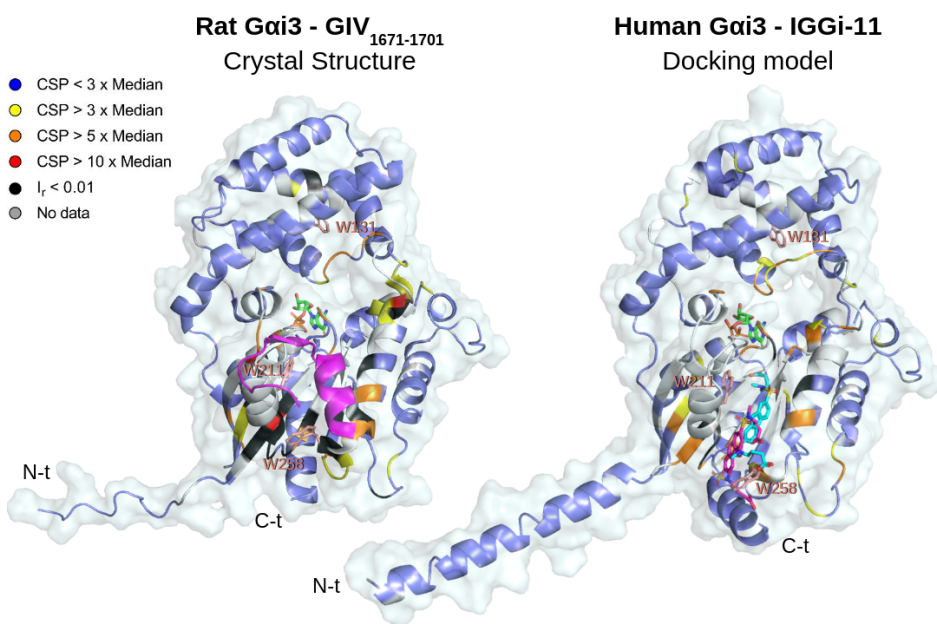
**Figure 17.** Plots of backbone amide NMR CSPs on GDP-Gai3 signals caused by GIV<sub>1671-1705</sub> (at 1:4 molar ratio) and IGGi-11 (1:6.4 molar ratio) for each residue. At the top, the location of helices,  $\beta$ -strands and switch regions are indicated. At the bottom, the CSPs are mapped onto a model structure of Gai3, generated from NMR chemical shifts using the web server CS23D (<http://www.cs23d.ca>), with GDP manually docked based on the crystal structure of GDP-bound Gai1. The protein transparent surface is pink for the helical domain and blue for the G-domain. A portion of the structure is zoomed in without CSP-based colors and showing the switch regions in gold.

To solve these ambiguities, we prepared samples with specific  $^{15}\text{N}$ -enrichment in the tryptophan residues of Gai3 and its W258F mutant. This mutant binds GIV<sub>1671-1701</sub> with a 4-fold reduced affinity (de Opakua et al., 2017). By comparing the NMR spectra of wild-type and the W258F mutant Gai3 bound to GIV<sub>1671-1705</sub> or IGGi-11, we were able to assign the signals of the three tryptophan side chains in the complexes. W211 (switch II region) is strongly perturbed, W258 ( $\alpha 3$ - $\beta 5$  loop) is perturbed, and W131 (helical domain) is unperturbed (**Figure 18**). These results indicate that those two regions of the G-domain are involved in binding, and that little or no long-range conformational changes occur in the helical domain.



**Figure 18.** Comparative spectral analysis of indole signals in selectively  $^{15}\text{N}$ -Trp labeled Gai3 W258F upon interaction with GIV<sub>1671-1705</sub> and IGGi-11. The spectrum in the top panel shows the effect on W131 and W211 side chain signals. In the middle and bottom panels the assignment of the indole signals in uniformly labeled Gai3 is shown in both complexes. The signal of W211sc is very weak in the NMR spectrum of the Gai3-IGGi-11 complex, and its location is indicated by "X" based on the  $^{15}\text{N}$ -Trp Gai3 W258F spectra of the top panel, which is consistent with the tracking of the signal along a titration.

These observations are consistent with the crystal structure of the rat GDP-Gai3 homolog bound to GIV<sub>1671-1701</sub> fragment GIV fragment (Kalogiropoulos et al, 2019) (**Figure 19**). Rat and human Gai3 have a sequence identity of 98.6%. The electron density shows most of the Gai3 residues (from D26 to N347) but only about half of GIV<sub>1671-1701</sub> (from G1673 to N1690). The structure shows GIV residues <sup>1682</sup>LQQFLEESN<sup>1690</sup> as an amphipathic helix bound to the hydrophobic pocket lined by switch II (where W211 is located), helix  $\alpha$ 3 and the N-terminal region of the  $\alpha$ 3- $\beta$ 5 loop (where W258 is located). The apolar side chains of GIV are buried in the cleft while the polar ones are solvent exposed. GIV residues <sup>1673</sup>GSPGSEVVT<sup>1681</sup> (folded into a  $\beta$ -strand followed by a coil structure) interact with switch II residues (folded into a  $\beta$ -strand and a helix) on the surface of Gai3.

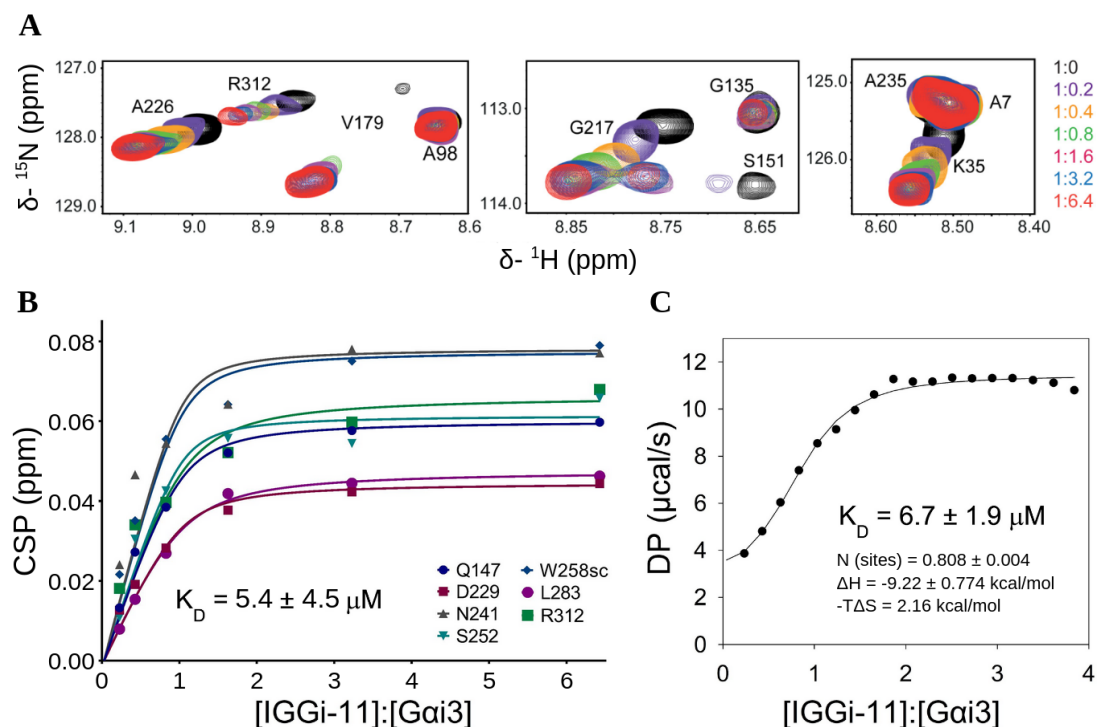


**Figure 19.** *Left panel:* Crystal structure (PDB 6MHF) of rat GDP-Gai3 bound to GIV<sub>1671-1701</sub>. The ribbon representing the backbone of the protein is colored according to the CSPs measured on human GDP-Gai3 bound to GIV<sub>1671-1705</sub> (as in Figure 15), and the GIV peptide is in magenta. GDP (green) and the three tryptophan side chains (pink) are shown in sticks. *Right panel:* homology model of human GDP-Gai3 with four poses of docked IGGi-11 (in sticks colored magenta or cyan). The ribbon representing the backbone of the protein is colored according to the CSPs measured on human GDP-Gai3 bound to IGGi-11 (as in Figure 15). The side chains of the three tryptophan residues are shown as pink sticks. The chain termini of the two proteins are indicated. Rat Gai3 has a non-native N-terminal region that interacts with symmetry mates in the crystal lattice, and several residues at the C-terminal region are not seen in the electron density. The model of human Gai3 was based on the crystal structure of human Gai1 bound to a non-natural peptide (PDB 1Y3A). This crystal structure lacked residues at the chain termini, and the native N-terminal and C-terminal regions of Gai3 were modeled as helices.



Our collaborators conducted a docking of IGGi-11 on a homology model of GDP-Gai3 (Zhao et al., 2023), which led to the selection of two poses (**Figure 19**). This model is consistent with experimental data (from NMR and mutagenesis experiments), and indicates that IGGi-11 binds to the same cleft on human Gai3 as GIV residues 1682-1690 do on rat Gai3 (the cyan IGGi-11 and GIV helix in **Figure 19**).

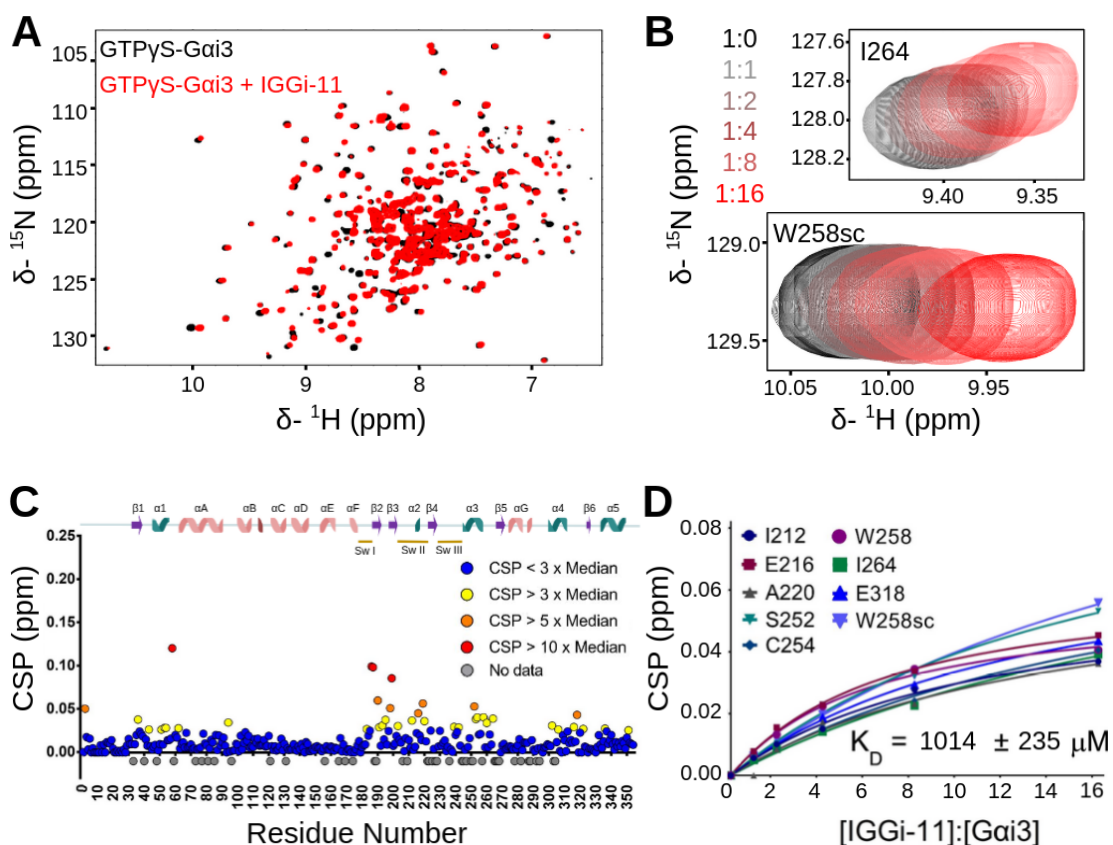
Along the titration of GDP-Gai3 with IGGi-11 the NMR signal of several residues exhibited gradual shifts that indicated an exchange kinetics in the intermediate-fast regime. The shifts of seven such signals could be fit to a binding equilibrium with assuming a 1:1 stoichiometry yielding a dissociation constant  $K_D = 5.4 \pm 4.5 \mu\text{M}$  at 30 °C (**Figure 20, B**). A similar result was obtained by ITC, with a  $K_D = 6.7 \pm 1.9 \mu\text{M}$  at 25 °C. The number of sites was an adjustable parameter in the fitting the ITC, and a value close to 1 (with  $N = 0.81$ ) indicates that the equimolar stoichiometry model is correct (**Figure 20, C**).



**Figure 20.** Affinity determination of the interaction of IGGi-11 with GDP-Gai3. (A) Illustration of the titration of Gai3 with IGGi-11 at 30 °C. (B) CSP measured along the NMR titration for seven selected residues of Gai3. The symbols are the experimental measurements and the line is the fitting. The indicated  $K_D$  value is the average  $\pm$  the standard deviation of the seven selected residues. (C) ITC experiment at 25 °C.

## 4.1.3 IGGi-11 specifically targets Gai3 in its GDP-bound state

NMR shows that IGGi-11 interacts very weakly with GTP-Gai3. The perturbations caused by a 16-fold excess of IGGi-11 on Gai3 bound to GTP $\gamma$ S (a non-hydrolysable analog of GTP) are small (**Figure 21 A**), and occur in regions that are also perturbed in GDP-Gai3 (**Figure 21 C**), indicating that the binding site is the same but with a lower affinity. The  $K_D$  value estimated from nine selected signals that shift along the titration is  $1014 \pm 235 \mu\text{M}$  at 30 °C (**Figure 21 B, D**). The estimated  $K_D$  has a large error because the titration is far from reaching saturation, but it shows that the affinity is about 200-fold lower than for GDP-Gai3.



**Figure 21.** (A) NMR spectrum displaying the overlay of GTP $\gamma$ S-Gai3 without (black) or with (red) IGGi-11 at molar ratio 1:16 (red). (B) Example of selected residues used for affinity determination, illustrating their chemical shift changes upon interaction with IGGi-11. (C) CSPs of backbone amide NMR signals in GTP $\gamma$ S-Gai3 induced by IGGi-11 at a 1:16 molar ratio for each residue. (D) CSP measured along the NMR titration for nine selected residues of Gai3. The symbols are the experimental measurements and the line is the fitting. The indicated  $K_D$  value is the average  $\pm$  the standard deviation of the nine residues.

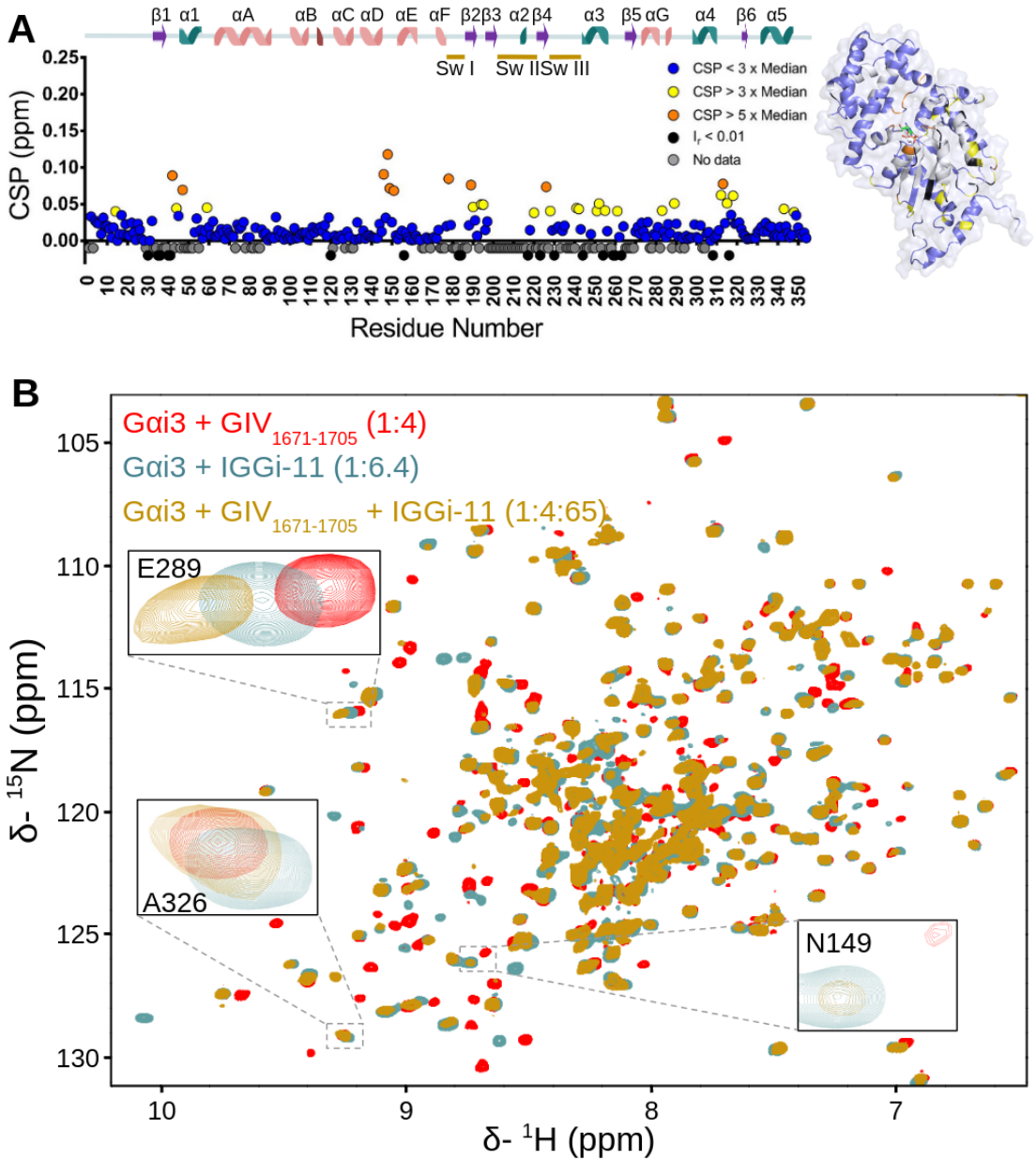
#### 4.1.4 IGGi-11's impact on Gai3-GIV<sub>1671-1705</sub> interaction

The competitive inhibition of the GDP-Gai3 interaction with GIV by IGGi-11 suggests that a sufficient excess of IGGi-11 should displace GIV from GDP-Gai3. We tested this on the NMR sample of GDP-Gai3 at the end of the titration with GIV<sub>1671-1705</sub> (with a 1:4 molar excess of the peptide). Since the affinity of GDP-Gai3 is approximately 10 times higher for the GIV peptide than for IGGi-11, a very large excess of IGGi-11 would be necessary, but practical limitations led us to add only a 65-fold molar excess of IGGi-11 (16-fold molar excess over the GIV peptide). When comparing the spectra of GDP-Gai3 in the presence of GIV<sub>1671-1705</sub>, IGGi-11 or both it is challenging to determine which complex or complexes exist in the ternary mixture. While certain signals align with those of the Gai3-GIV<sub>1671-1705</sub> or Gai3-IGGi-11 complexes, there are also unique signals present in the mixture (**Figure 22**).

We considered four possible scenarios:

1. Complete displacement: The NMR spectrum of the ternary mixture should be identical to that of Gai3-IGGi11's spectrum. Thus, when using Gai3-IGGi11 as the reference for CSP calculation, the average CSP should be zero.
2. No displacement: The NMR spectrum of the ternary mixture should be identical to that of Gai3-GIV<sub>1671-1705</sub>'s spectrum. In this case, when using Gai3-GIV<sub>1671-1705</sub> as the reference for CSP calculation, the average CSP should be zero.
3. Coexistence of complexes: If both Gai3-GIV<sub>1671-1705</sub> and Gai3-IGGi-11 complexes coexist in the sample, the NMR spectrum of the sample containing Gai3-GIV<sub>1671-1705</sub>-IGGi-11 should exhibit characteristics of both NMR spectra.
4. Formation of a new ternary complex: If a new ternary complex is formed, it would result in a distinct NMR spectrum that differs from both the Gai3-IGGi-11 and Gai3-GIV<sub>1671-1705</sub> spectra.

Calculating the CSPs measured in the ternary mixture using the NMR spectrum of Gai3-GIV<sub>1671-1705</sub> as a reference yields an average CSP of  $0.020 \pm 0.028$  ppm. Similarly, using the NMR spectrum of Gai3-IGGi11 as a reference, results in an average CSP of  $0.015 \pm 0.014$  ppm.



**Figure 22.** (A) CSPs of backbone amide NMR signals in Gai3-GIV induced by IGGi-1 at a 1:4:65 molar ratio. The CSPs are mapped on the structural model of Gai3, generated from backbone chemical shifts and with a docked GDP. (B) Overlay of NMR spectra of the GDP-Gai3 with GIV1671-1705 (1:4 molar ratio), IGGi-11 (1:6.4 molar ratio) complex, or both (with IGGi-1 at a 1:65 molar ratio). The zoomed-in regions reveal that the spectrum of the mixture contains residues resembling Gai3-GIV (A326), residues resembling Gai3-IGGi-11 (N149), and residues that do not resemble either of the two spectra (E289). Consequently, it is challenging to draw a conclusion about which spectrum the overall mixture resembles.

---

A smaller average CSP indicates a greater resemblance to the corresponding reference complex. However, in this case, the average CSP values are within the statistical uncertainty, making it not possible to draw a definitive conclusion. Probably a ternary complex is present. The CSPs induced by IGGi-11 on G $\alpha$ i3 indicate that IGGi-11 binds within the G-domain although the binding site of IGGi-11 is uncertain. It is possible that the hydrophobic IGGi-11 molecule still has some affinity for part of the hydrophobic cleft between switch II and  $\alpha$ 3 when the GIV peptide is bound.

#### **4.1.5 IGGi-11 hinders signaling mediated by the GIV-G $\alpha$ i3 interaction**

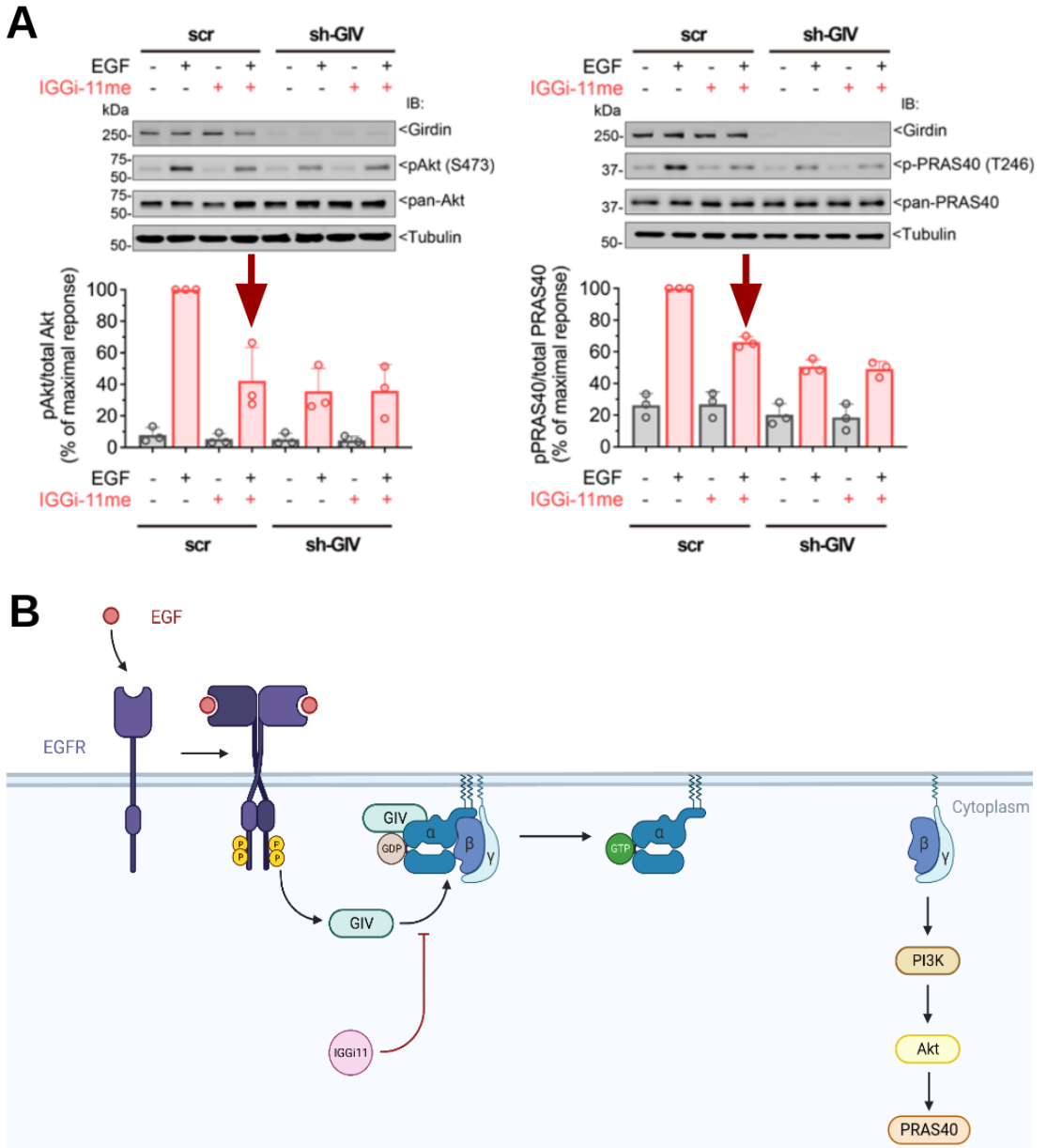
Upon stimulation by growth factors, the interaction of EGFR with GIV is reinforced, a ternary complex with G $\alpha$ i3 is assembled, and signaling cascades are triggered by GTP-G $\alpha$ i3 and by the dissociated G $\beta\gamma$ , with an impact in cell proliferation and migration. IGGi-11 disrupts the GIV-G $\alpha$ i3 interaction and reduces the migration phenotype of MDA-MB-231 cells, from a metastatic breast cancer cell line with high GIV expression levels.

To explore if the effect of IGGi-11 could be acting on signaling cascades other than those initiated by EGFR and mediated by the GIV interaction with GDP-G $\alpha$ i3, we measured the phosphorylation levels of downstream targets of the cascade in MDA-MB-231 cells under the following conditions:

- Stimulation with EGF.
- Stimulation with EGF and treatment with IGGi-11me.
- GIV-depletion (using shRNA), and stimulation with EGF.
- GIV-depletion (using shRNA), stimulation with EGF, and treatment with IGGi-11me.

The hypothesis is that if IGGi-11 treatment only inhibits the G $\alpha$ i3-GIV interaction, the result obtained when treating stimulated cells should be similar to the result with stimulated GIV-depleted cells, both treated and untreated with IGGi-11me. Differences in the measurements could indicate that IGGi-11 is altering other pathways.

In these assays, the methyl-ester form of IGGi-11 was used (IGGi-11me). This modification enhances its cellular permeability, and intracellular esterases produce IGGi-11. Therefore, IGGi-11me is the prodrug of the IGGi-11 drug.



**Figure 23.** (A) Western blot analysis demonstrating the inhibitory effect of IGGi-11 on the Gai3-GIV interaction in stimulated MDA-MB-231 cells. The phosphorylation levels of Akt and its substrate, PRAS40, are reduced in cells treated with IGGi-11me (red arrow) compared to untreated cells, resembling the levels observed in GIV-deficient MDA-MB-231 cells. These results indicate successful inhibition of the Gai3-GIV interaction. "scr" denotes control cells with a scrambled short hairpin RNA, while "sh-GIV" represents GIV-depleted cells (small hairpin RNA that blocks GIV expression). (B) Proposed mechanism of action of IGGi-11 in the inhibition of the Gai3-GIV interaction.

Treatment with IGGi-11me decreases Akt phosphorylation and its substrate PRAS40, without changes in the amount of Akt or PRAS40 proteins, as evidenced by the Western blot analysis (**Figure 23, A**). These results strongly suggest that IGGi-11 specifically disrupts the signaling cascades triggered by the G $\alpha$ i3-GIV interaction. These findings underscore the potential as a therapeutic agent for modulating these important cellular processes (**Figure 23, B**).

## 4.2 Crystal structure of Gai3

While the NMR chemical shift mapping provides structural information on the Gai3-IGGi-11, higher resolution information would be very useful to better understand the basis of the molecular recognition and, eventually, conduct structure-activity-relationship studies using modified compounds based on IGGi-11. Attempts to crystallize GDP-Gai3 using the 1-354 sequence preceded by the non-native five residue long N-terminal tail were unsuccessful, as were attempts using a 31-354 sequence (with the same five residue long non-native N-terminal tail). This N-terminal deletion mutant had been studied by NMR, showing a similar behavior as the full-length Gai3 sequence. The G $\alpha$  subunits have an N-terminal region of about 35 residues that interacts with the G $\beta$  in the trimeric form of the G protein. When G $\alpha$  subunits are dissociated from G $\beta\gamma$ , the N-terminal region is believed to be disordered. Because it contains the sites of fatty-acid attachment for lipid membrane anchoring (in Gai3, residue C3 is S-palmitoylated and G2 is N-myristoylated, after the removal of M1) it might have some conformational restrictions. But in soluble G $\alpha$  subunits produced in bacterial cultures for structural studies (without any post-translational modification) the N-terminal region appears to be dispensable for structural integrity and ligand binding, at least for some of them. It is likely that the disordered regions make it difficult to crystallize Gai3. There are 5 crystal structures of GDP-Gai3, in all of them bound to other regulatory proteins (or protein fragments), suggesting that the ligands and/or the conformational changes (or changes in the protein dynamics) induced by them facilitate crystallization. In two of them most of the N-terminal region is present but not visible in the electron density. Therefore, a shortened version of Gai3, deleting the most flexible terminal regions, could be more favorable for co-crystallization of the protein bound to IGGi-11.

The rat GDP-Gai3 homolog has been crystallized bound to a GIV peptide (Kalogiropoulos et al., 2019). An N-terminal deletion mutant did not produce crystals, but substituting the N-terminal 25 residues by a non-native sequence (consisting of a poly-His affinity purification tag and a linker) did, showing part of the non-native sequence ordered and making intramolecular contacts with a crystal symmetry mate. The C-terminal 7 residues of Gai3 were not visible in the crystal (indicating disorder). The co-crystallized 31-residue long human GIV<sub>1671-1701</sub> fragment showed that GIV residues <sup>1682</sup>LQQFLEESN<sup>1690</sup> form an amphipathic helix bound to the hydrophobic cleft lined by switch II, helix  $\alpha$ 3 and the N-terminal region of the



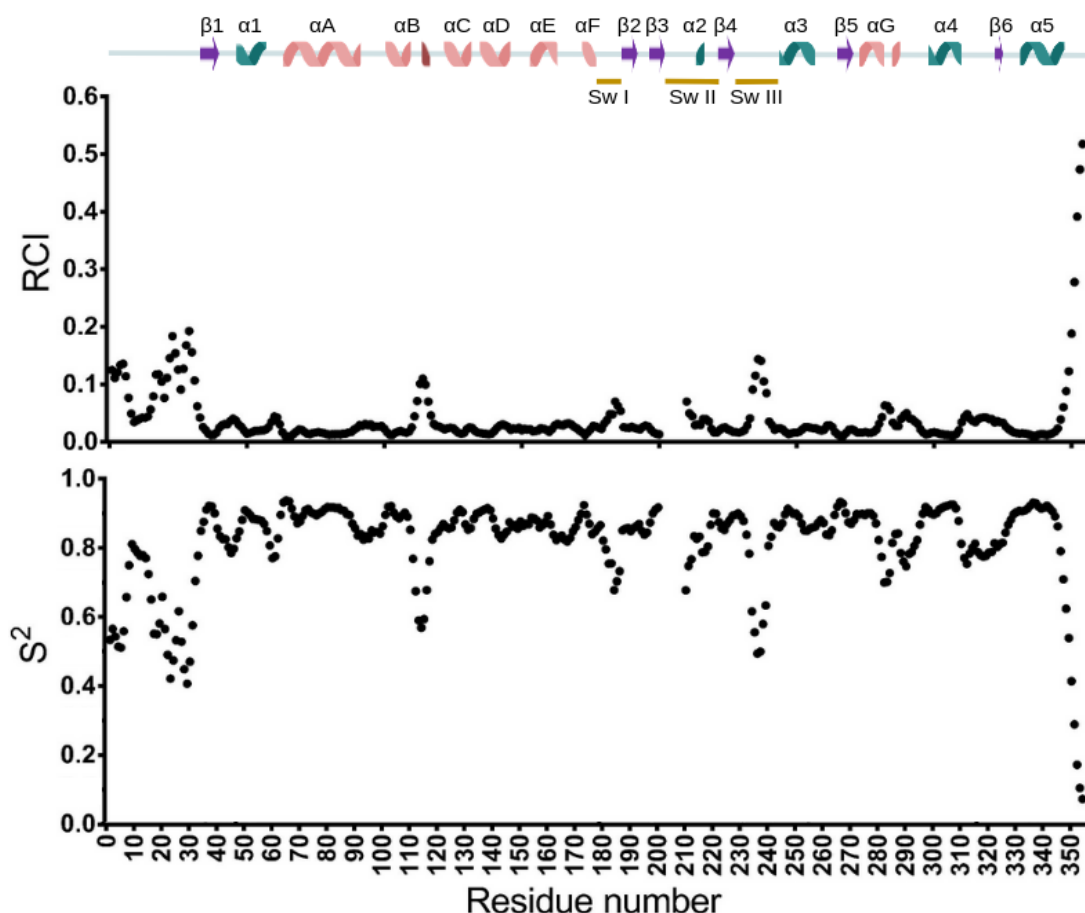
$\alpha$ 3- $\beta$ 5 loop of Gai3. The hydrophobic side chains on the GIV helix are buried while the hydrophilic ones are solvent exposed. GIV residues <sup>1673</sup>GSPGSEVVT<sup>1681</sup> (folded into a  $\beta$ -strand followed by a coil structure) interact with switch II residues (folded into a  $\beta$ -strand and a helix) on the surface of Gai3. It is likely that both the crystal contacts of the N-terminal non-native residues and a conformational change or reduced dynamics caused by GIV binding facilitated co-crystallization of rat GDP-Gai3 with the GIV peptide.

In that crystal, only 18 of the GIV<sub>1671-1701</sub> residues were visible, suggesting that a shorter peptide could be used for crystallization trials with a human Gai3 and, perhaps, further facilitate crystallization. A peptide corresponding to GIV<sub>1671-1692</sub> should bind with the same affinity. This peptide would, as happens with other about 30-residue long GIV fragments containing the GBA motif, have low solubility in aqueous buffer and no aromatic residues. These experimentally unfavorable properties are typically overcome by preparing peptide stocks in DMSO and by measuring the peptide concentration by absorbance at 205 nm. To use a peptide that could be prepared at high concentration in the same buffer as Gai3, we added an aspartic residue at the C-terminus end (increasing the net negative charge of the peptide at neutral pH). To measure the peptide stock concentration with higher accuracy by absorbance at 280 nm, a tyrosine residue was added after the aspartic one. This design yielded peptide GIV<sub>1671-1692-DY</sub> as a candidate for co-crystallization with a shortened version of human GDP-Gai3.

#### 4.2.1 Experimental mapping of flexible regions in GDP-Gai3

The NMR assignment of human Gai3 allows us to make a prediction on the flexibility of the backbone at the residue level (Berjanskii & Wishart, 2005). The higher the value of the so-called Random Coil Index for a given residue the higher the flexibility of the backbone at that residue. For GDP-Gai3, high flexibility is predicted for most of the N-terminal 30 residues and the C-terminal 8 residues (**Figure 24**). High flexibility is also predicted for the C-terminal half of the second helix of the helical domain and for the switch III region of the G-domain. The RCI prediction algorithm is based on a statistical analysis of protein NMR chemical shifts and molecular dynamics simulations to quantify disorder with an arbitrarily scaled number. This number shows a correlation with the square of the so-called generalized order parameter (S) of the <sup>1</sup>H-<sup>15</sup>N bond, which is a measure of the spatial restriction

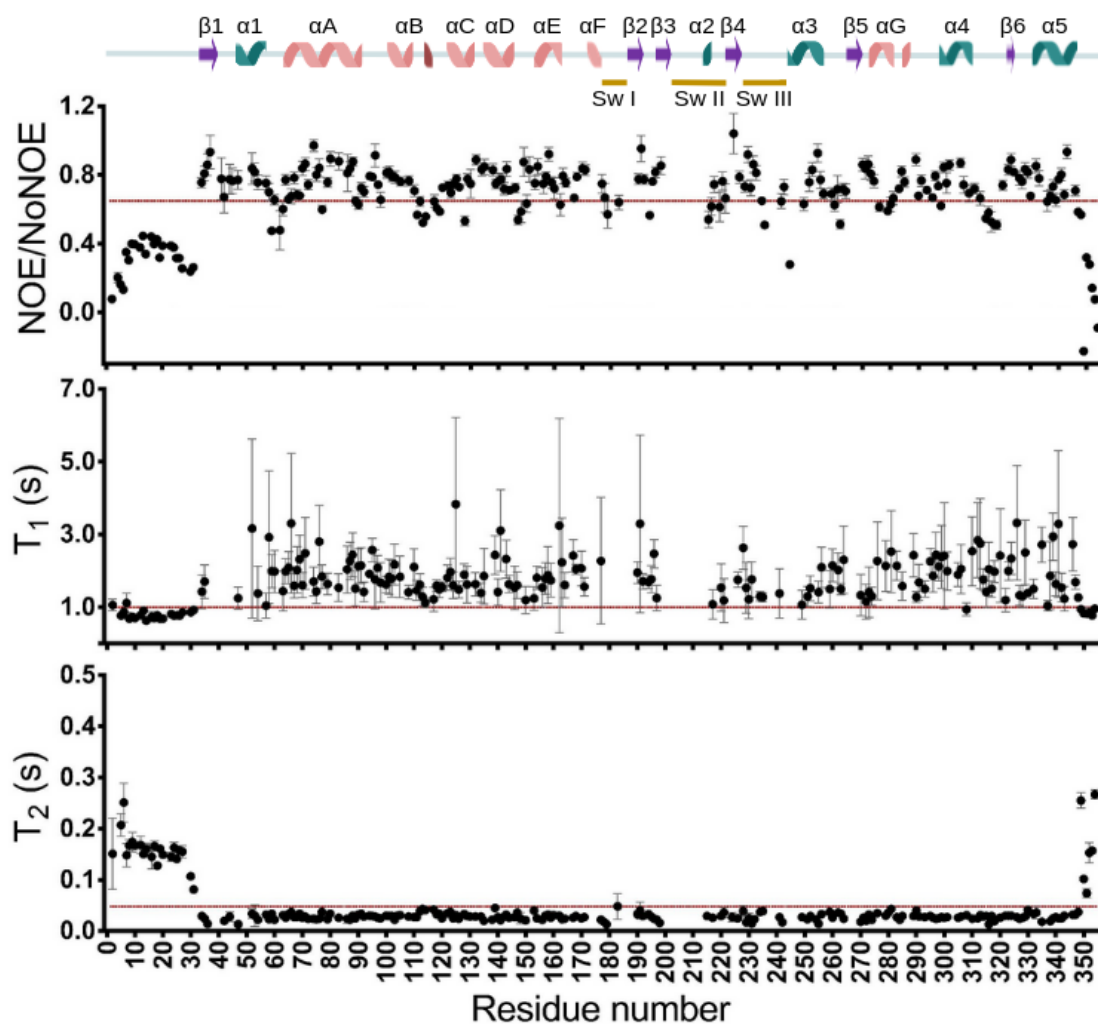
of its motion in the ps-ns time scale.  $S^2$ , can take values between 0 (no restriction, full disorder) and 1 (complete restriction, full order). The values of  $S^2$  estimated from the RCI are also represented in **Figure 24**.



**Figure 24.** Prediction of protein backbone flexibility in GDP-Gai3. The Random Coil Index (RCI) and  $S^2$  parameters were calculated from the assigned NMR chemical shifts using the RCI web server ([http://www.randomcoilindex.ca/cgi-bin/rci.cgi\\_current.py](http://www.randomcoilindex.ca/cgi-bin/rci.cgi_current.py)). The values for N-terminal non-native residues are colored grey. At the top, the location of helices,  $\beta$ -strands and switch regions are indicated.

To experimentally assess the degree of disorder along the backbone of GDP-Gai3 we measured three  $^{15}\text{N}$ -NMR relaxation parameters that inform about the local dynamics of the backbone on the ps-ns time scale. These parameters ( $T_1$ ,  $T_2$ , and the  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE) can be used to obtain the order parameter  $S^2$  by fitting them to an equation derived from a mathematical formalism that describes the local dynamics separately

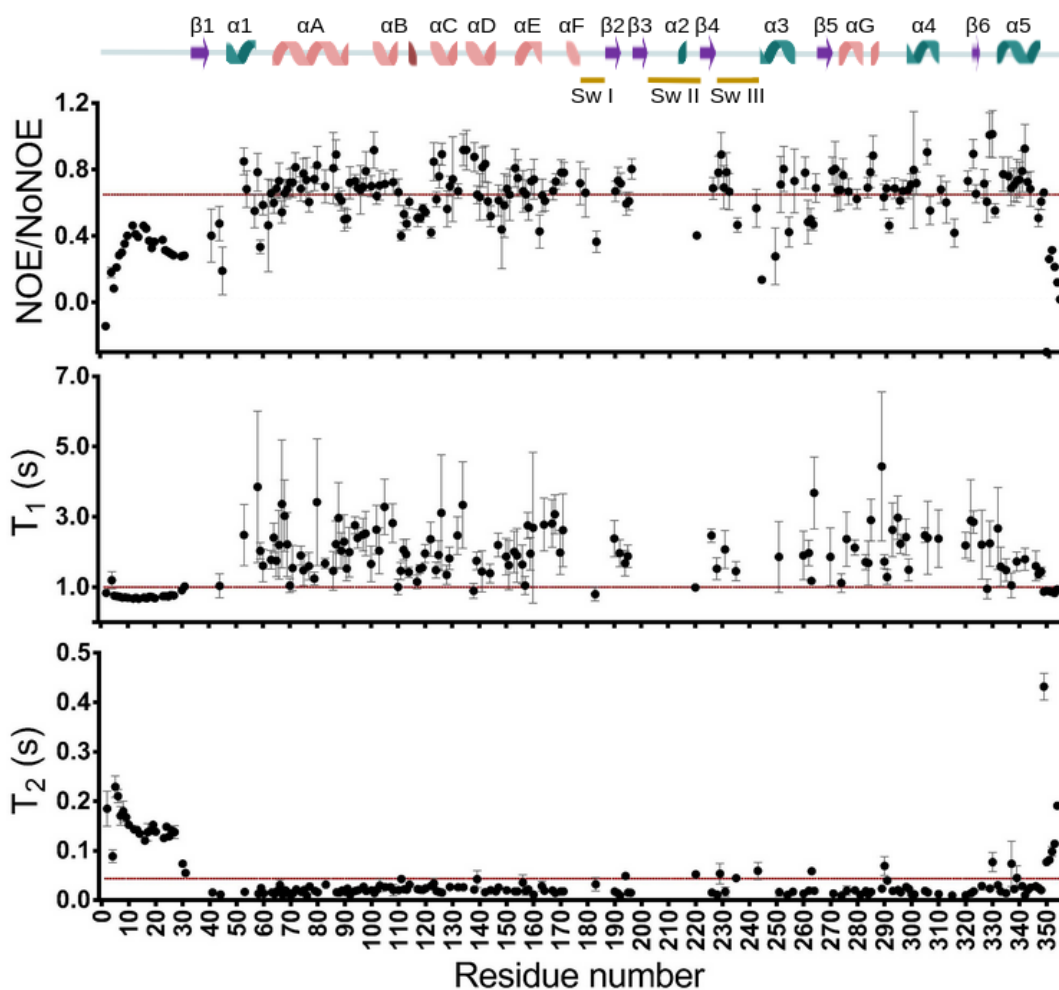
from the global dynamics (overall tumbling of the protein in solution). The values of these parameters that could be measured are shown in **Figure 25**.



**Figure 25.**  $^{15}\text{N}$ -NMR relaxation parameters measured on GDP-Gai3. The plot depicts the measured values of the heteronuclear NOE,  $T_1$ , and  $T_2$  along the protein sequence. The red dotted line represents the threshold used to identify the most flexible regions.

The fitting of the data to calculate the order parameters was not successful. For many of the residues the value of  $S^2$  was almost 1.0 with a very small error (for instance,  $S^2 = 0.99 \pm 0.01$  for A300), and an unusually large range of effective correlation times, and with large errors, calculated along the sequence (with values from a few ps to 2 ns). Therefore, this analysis was discarded.

Although some of the measurements have large errors (especially  $T_2$  values) they are reliable enough to delineate the most mobile regions. Both the heteronuclear NOE and  $T_2$ , show the N-terminal and C-terminal tails to be much more mobile than the rest of the chain. The  $T_1$  values also show this higher mobility of the chain ends, although the large dispersion of the values in the less mobile region of the proteins (and the large errors) indicate that these measurements are less reliable.



**Figure 26.**  $^{15}\text{N}$ -NMR relaxation parameters measured on GDP-Gai3 bound to GIV<sub>1660-1870</sub>. The plot depicts the measured values of the heteronuclear NOE,  $T_1$ , and  $T_2$  along the protein sequence. The red dotted line represents the threshold used to identify the most flexible regions.

With a threshold of 0.65, the heteronuclear NOE defines the first 31 residues in the N-terminal region and the last 8 residues in the C-terminal region as the most flexible regions.  $T_2$  relaxation times above the mean value (0.048), define as most mobile the

first 31 residues and the last 6 residues.  $T_1$  relaxation times below 1.0 define the first 31 residues and the last 6 residues as the most flexible regions. Thus, the first 31 residues in the N-terminal region (no measurement could be obtained for K32 and E33), and the 6 last residues in the C-terminal region consistently exhibited very high relative flexibility. GDP-Gai3 bound to GIV<sub>1660-1870</sub> showed similar results (**Figure 26**).

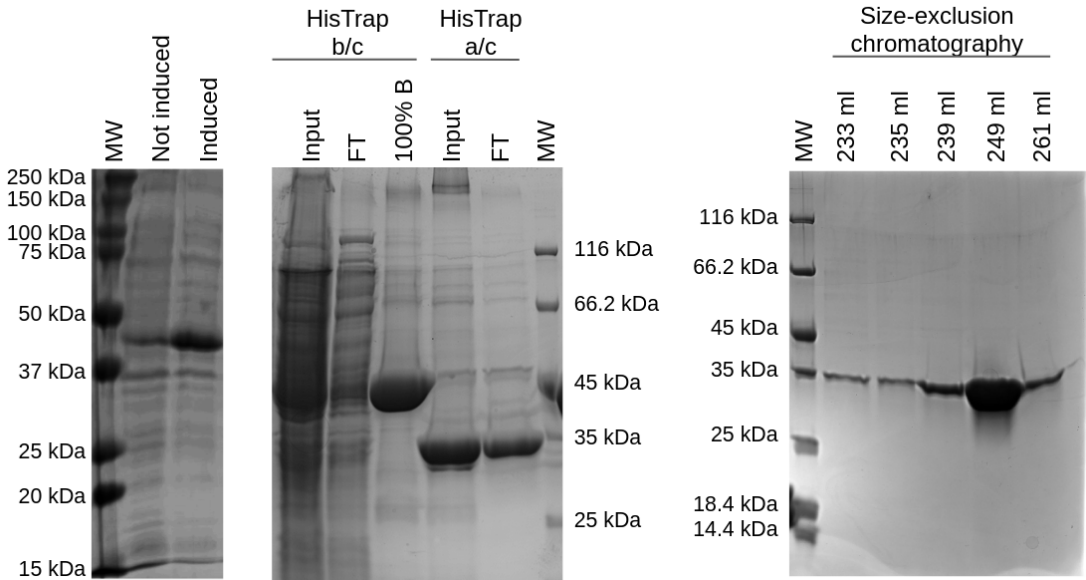
We hypothesized that deleting the terminal flexible tails would yield a shortened version of Gai3 with the same ligand binding properties but easier to crystallize than the full-length one, and designed a gene to produce Gai3<sub>31-350</sub> (hereafter named sGai3) by removing the first 30 residues and the last 4 residues. The clone for protein production included a protease site that, after cleavage, will leave a non-native Gly residue preceding the first native one (A31), thus including at least two flexible residues before the one whose flexibility could not be assessed (K32). Similarly, at the C-terminus four non-native residues were deleted, leaving two flexible ones (K349 and E350), which might also favor solubility (considering the preceding hydrophobic residue).

#### 4.2.2 Production and purification of sGai3

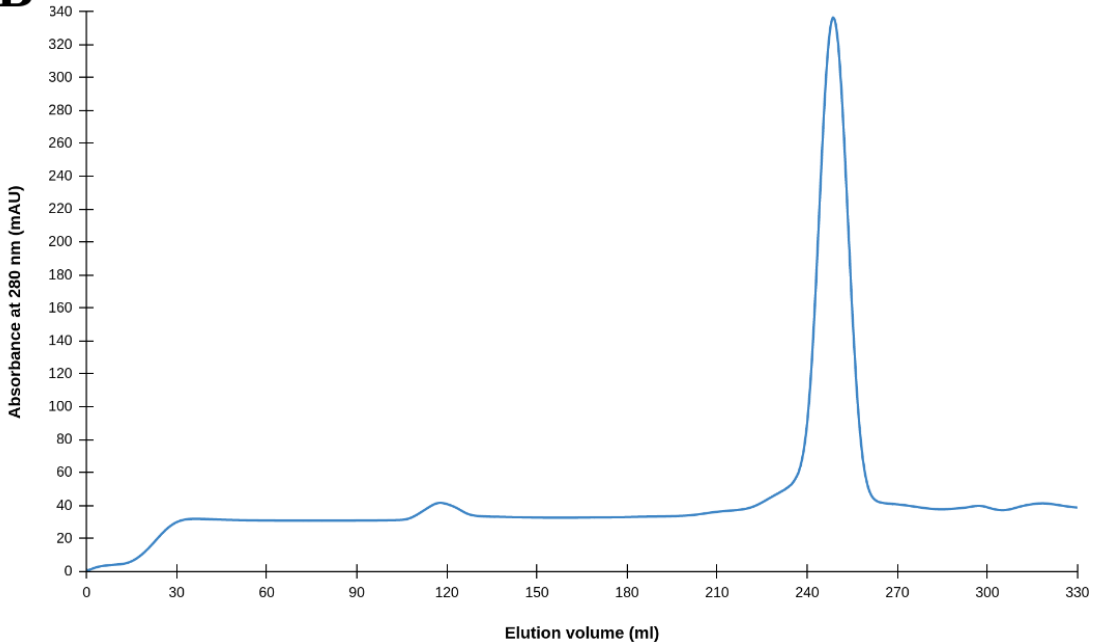
The gene utilized for the production of sGai3 was designed and synthesized with codons optimized for *E. coli* and cloned in plasmid pET29a(+). It consisted of an N-terminal fusion with ubiquitin with a polyHis sequence for affinity purification, a short linker, a TEV cleavage site, and Gai3<sub>31-350</sub>. The fusion with ubiquitin has been a common strategy in the laboratory to enhance expression levels and improve solubility of the target protein.

To express sGai3, a three-day growth process was conducted. Initially, a preinoculum was prepared using freshly transformed BL21(DE3) cells, which were incubated in LB medium with kanamycin at 37 °C overnight. Subsequently, 50 ml of the preinoculum was added to each of the Erlenmeyer flasks containing 0.5 l of in peptone-yeast extract-NaCl medium with kanamycin. The cultures were incubated at 37 °C until they achieved an optical density of 0.8 at 600 nm. After reaching this threshold, IPTG was added at a concentration of 1 mM to induce gene expression, followed by 16-20 h of overnight incubation at 20 °C. The cells were then centrifuged at 3,500 rpm for 20 minutes to pellet them, and the resulting pellet was resuspended in the lysis buffer for further processing.

**A**



**B**

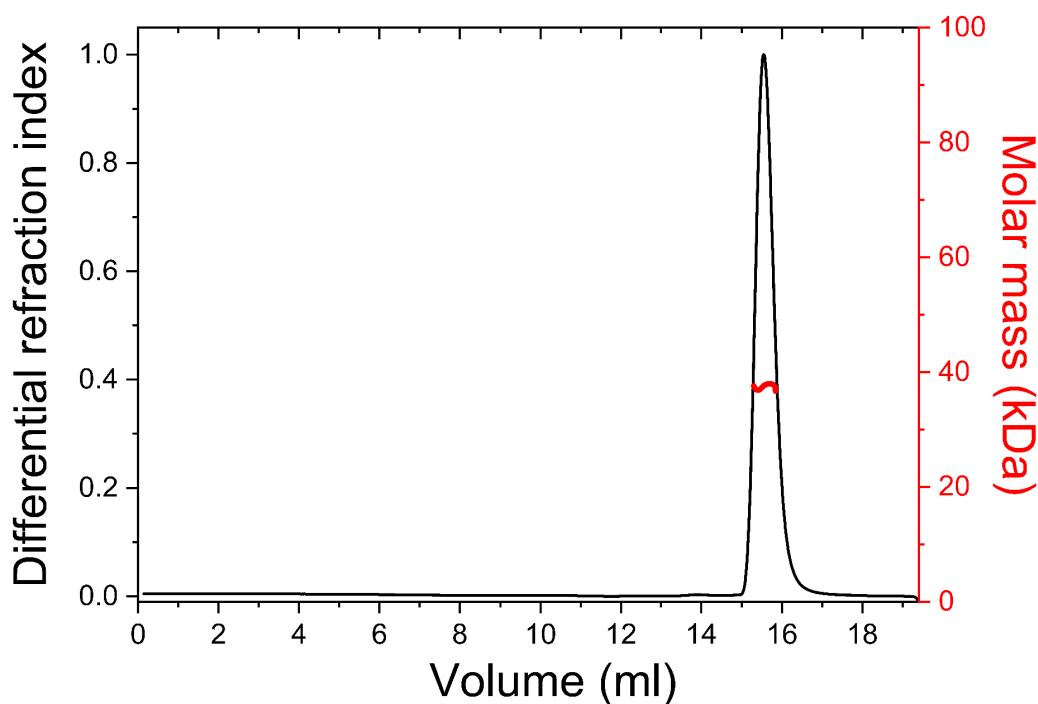


**Figure 27.** (A) Summary of the chromatographic purification of sGai3 monitored by reducing SDS-PAGE. *Left:* Cultures pre- and post-induction. *Middle:* Input, flowthrough (FT) and elution (100 %B) of Ni<sup>2+</sup> chromatography before (b/c) and after (a/c)TEV cleavage. *Right:* Fractions from the size exclusion chromatography. (B) Elution profile of the size exclusion chromatography on a Superdex 200 26/60 column.

The initial step of the purification involved the use of affinity chromatography, which successfully eliminated most of the contaminants. The cleavage of His-sGai3 with TEV protease (overnight incubation at 4 °C during dialysis to remove imidazole), was highly efficient, resulting in almost complete cleavage, with sGai3 recovered in the flow-through of a second affinity chromatography. A final purification step by size exclusion chromatography revealed a highly pure sGai3 eluting as a symmetric peak at the expected volume for its monomeric molecular mass (**Figure 27**). All purification steps, except TEV cleavage, were done at room temperature. The yield was about 7 mg of pure sGai3 protein per liter of culture. The same procedure to produce and purify  $^2\text{H}$ - $^{15}\text{N}$  isotopically enriched protein (from bacteria cells grown in isotopically supplemented media) yielded 2 mg of pure protein per liter of culture.

#### 4.2.3 Suitability of sGai3 construct for crystallization assays

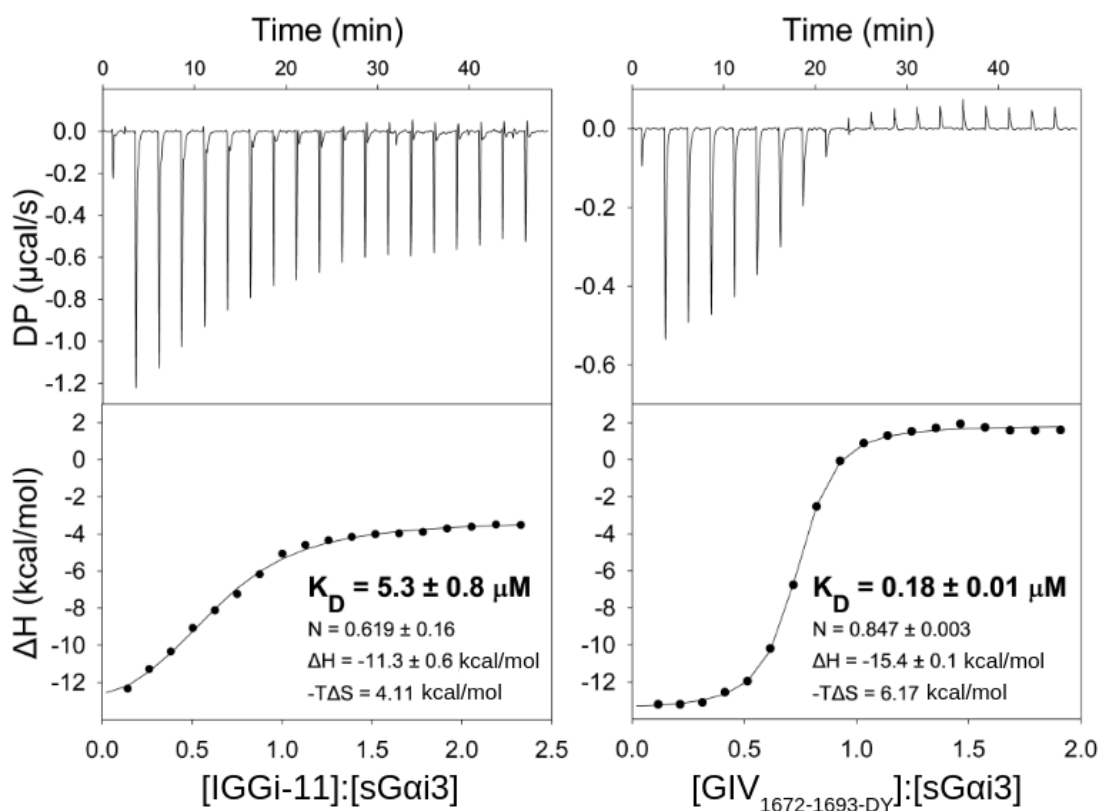
A SEC-MALS analysis yielded a single peak with a molar mass of 37.5 kDa (**Figure 28**), very close to the 36.8 kDa calculated from the amino acid sequence. This result confirmed the monomeric nature of sGai3.



**Figure 28.** SEC-MALS analysis of sGai3, showing the elution profile at 23 °C in PBS, pH 7.4. This single peak confirms the monomeric nature of the sample, with a molar mass corresponding to 37.5 kDa, consistent with the expected value for sGai3.

The binding of IGGi-11 and GIV<sub>1672-1693-DY</sub> to sGai3 was assessed by isothermal titration calorimetry (ITC) to determine if the shorter protein exhibited the same behavior as full-length Gai3 (Figure 29).

The results show that IGGi-11 binds sGai3 with the same affinity as full-length Gai3. They also show that the short GIV peptide with the GBA motif and two C-terminal non-native residues binds sGai3 with a very similar affinity as longer GIV peptides bind to full-length Gai3. Therefore, sGai3 and GIV<sub>1672-1693-DY</sub> are suitable molecules to characterize the molecular recognition events and potentially more favorable for co-crystallization trials.



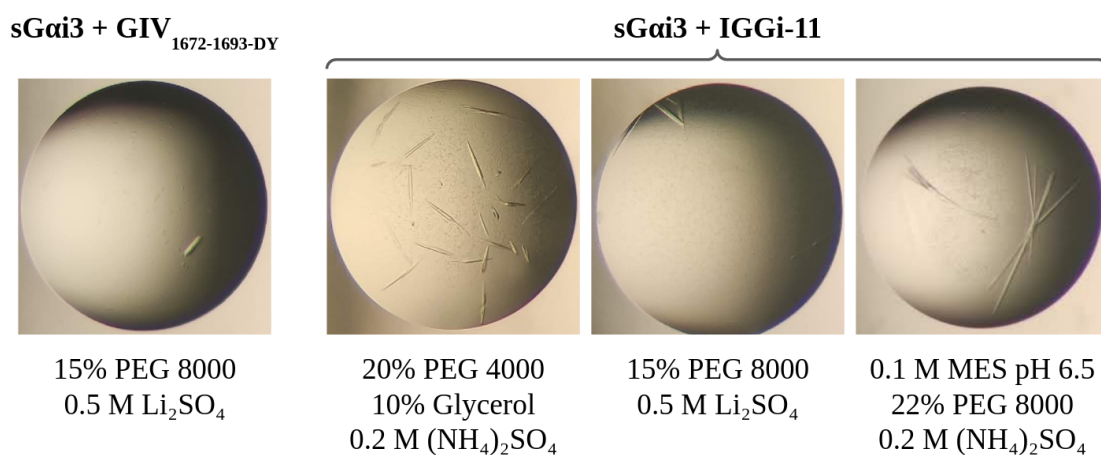
**Figure 29.** Isothermal Titration Calorimetry experiments at 25 °C to measure the binding of IGGi-11 and GIV<sub>1672-1693-DY</sub> to GDP-sGai3. The data, obtained at 25 °C, were fitted to a 1:1 stoichiometry binding model. The values of  $N$  close to 1 indicate correct stoichiometry, as observed for the GIV<sub>1672-1693-DY</sub> interaction with GDP-sGai3. However, for IGGi-11, the apparent low  $N$  value is likely due to the theoretical, unmeasured concentration of IGGi-11, suggesting that it may have been underestimated and thus aligning with a higher  $N$  value indicative of correct stoichiometry.



#### 4.2.4 Crystallization, data collection and refinement

Attempts were made to obtain crystals of sGai3 and of its complexes with GIV<sub>1672-1693-DY</sub> or IGGi-11. An initial screening was conducted using commercially available solutions using a stock of 10 g/l of sGai3 with or without a 1:2 molar ratio of one of the ligands.

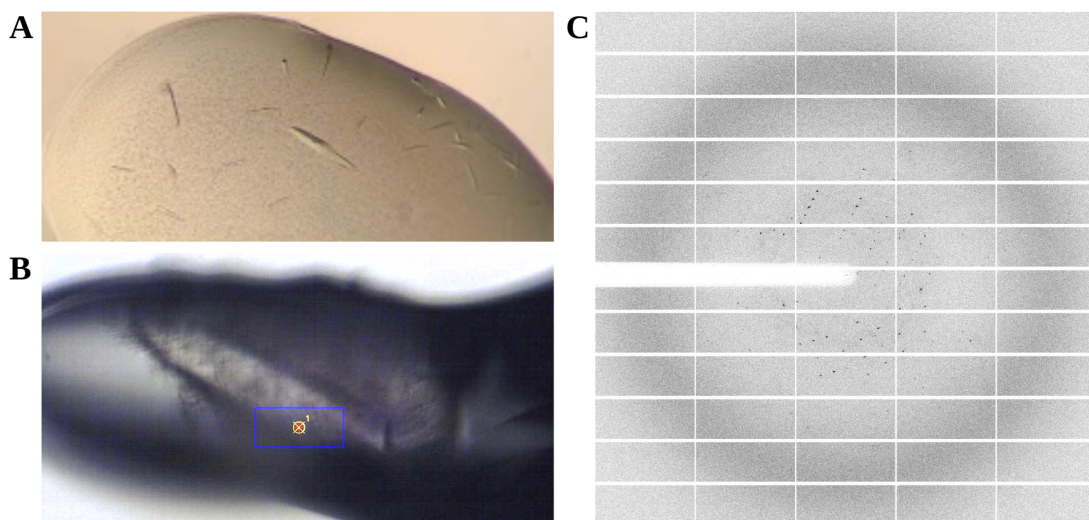
Protein crystals were successfully obtained only in drops containing either one of the two ligands, displaying a blade-like crystal habit. No crystals were obtained in drops with sGai3 alone, despite testing various conditions. As a result, all subsequent experiments were solely focused on the sGai3-GIV<sub>1672-1693-DY</sub> and sGai3-IGGi-11 complexes (**Figure 30**).



**Figure 30.** Photographs of drops from the initial screenings showing protein crystals in the presence of the ligands. Below the photograph, the differential additives in the mother liquor are indicated. No crystals were obtained for the free protein.

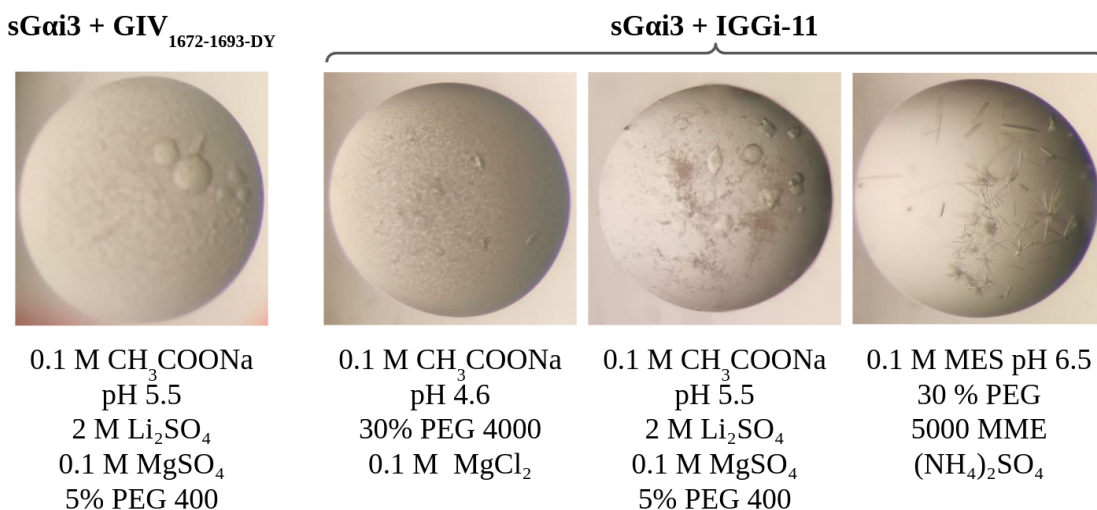
After several rounds of condition optimization, we successfully obtained crystals of sufficient size to be diffracted at the ALBA synchrotron. However, despite achieving diffraction from most of the crystals, the resolution obtained was approximately 7 Å. Thus, it was not possible to determine the structure.

Following a comprehensive detergent screening and subsequent optimization rounds, we obtained new large crystals that were cryo-protected crystals and rapidly frozen in liquid nitrogen for submission to the ALBA Synchrotron. While sGai3 with IGGi-11 yielded well-diffracting crystals (**Figure 31**), the crystals of sGai3 with GIV<sub>1672-1693-DY</sub> exhibited a resolution of 7-8 Å, which is insufficient for structure determination.



**Figure 31.** Crystal handling and diffraction pattern of GDP-sGai3 with IGGi-11. (A) Image showing the drop containing the collected crystals, in 0.1 M MES pH 6.5, 30% PEG 5000 MME, and 0.2 M ammonium sulfate. (B) Close-up view of the crystal mounted in the loop. (C) Diffraction pattern obtained from the crystal, revealing the scattering of X-rays and providing information about the crystal's structure.

Crystal diffraction analysis showed a space group  $P 4_2 2_1 2$ , with a single molecule in the asymmetric unit. Molecular replacement and structure refinement (Table 4) showed only GDP-sGai3 in the asymmetric unit, with no IGGi-11 molecule (**Figure 33**, see below).



**Figure 32.** Images showing the drops obtained during the initial screenings for both complexes using the crystals from the sGai3-IGGi-11 complex as seeds. Below the photograph, the differential additives in the mother liquor are indicated.

This finding prompted us to further pursue the crystallization of the complexes by using crystals as seeds for new crystallization assays. The addition of seeds resulted in crystals growing under different conditions and with different crystal habits (tear-shaped) (**Figure 32**), suggesting a potential change in the space group, which could have implications for the binding of ligands to sGai3.

Unfortunately, the crystals of sGai3-GIV<sub>1672-1693-DY</sub> once again did not provide sufficient diffraction quality for data collection (5-7 Å). However, we were able to collect data for crystals of GDP-sGai3 with IGGi-11, and the diffraction data revealed a different space group (P 4<sub>3</sub> 2 2) and two molecules in the asymmetric unit.

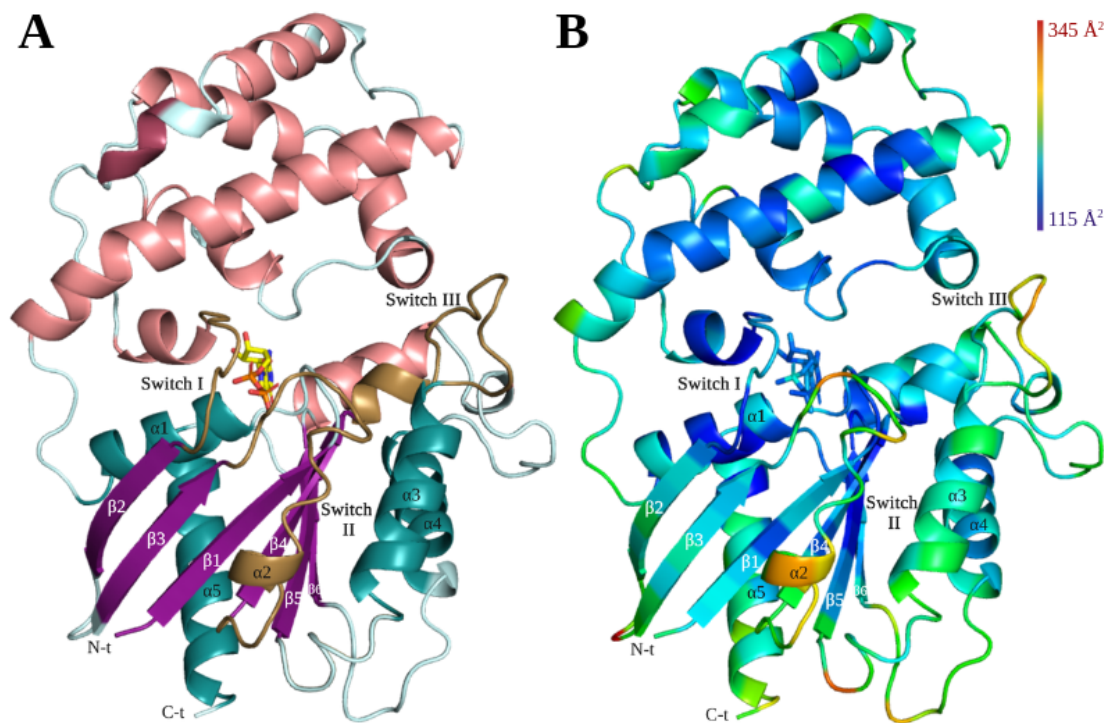
**Table 4.** Diffraction data collection and structure refinement statistics.

Data collection		Refinement	
Space group	P 4 <sub>2</sub> 2 <sub>1</sub> 2	Resolution (Å)	58.70 - 3.34
Wavelength (Å)	0.97926	No. reflections	5836
Cell dimensions		R <sub>work</sub> /R <sub>free</sub>	20.8 / 28.0
a, b, c (Å)	113.06, 113.06, 68.68	No. atoms (non-hydrogens)	
Resolution(Å) <sup>a</sup>	58.70-3.34 (3.52-3.34)	Protein (residues)	2544 (315)
Total reflections	109524 (5834)	GDP (molecules)	28 (1)
Unique reflections	5841 (292)	Average B factors (Å <sup>2</sup> )	
R <sub>merge</sub>	8.20 (4.22)	Protein atoms	166.75
R <sub>meas</sub>	8.40 (4.33)	GDP	137.88
R <sub>pim</sub>	2.00 (0.96)	R.m.s. deviations	
CC 1/2	99.9 (47.3)	Bond lengths (Å)	0.011
Completeness (%)	94.9 (79.0)	Bond angles (°)	1.300
<I/σ(I)>	21.6 (0.9)	Ramachandran statistics	
Wilson B factor	156.44	Preferred (%)	89.78
Multiplicity	18.8 (20.0)	Allowed (%)	7.67
Molecules per asymmetric unit	1	Outliers (%)	2.56
		Clashscore	19.03
		PDB code	8OY1

<sup>a</sup> Values in parentheses are for highest-resolution shell.

Interestingly, despite the structural differences observed compared to previous findings, we could not detect the presence of IGGi-11 in the crystal structure. The resolution of the data was lower than that of the previous analysis and the structure was not further refined.

The crystal structure, at 3.34 Å resolution, shows residues 33-347 of the protein (all of them except three at the N-terminus and three at the C-terminus). They are folded into a helical domain and a G-domain, with GDP bound in a pocket between them (**Figure 33**). The resolution is not sufficient to say if a Mg<sup>2+</sup> ion is present. The B-factors of the C $\alpha$  atoms (frequently interpreted as a measure of the atomic displacement due to local flexibility) are the largest for residues in the loops, and are large for switch II and III regions, but not for switch I.



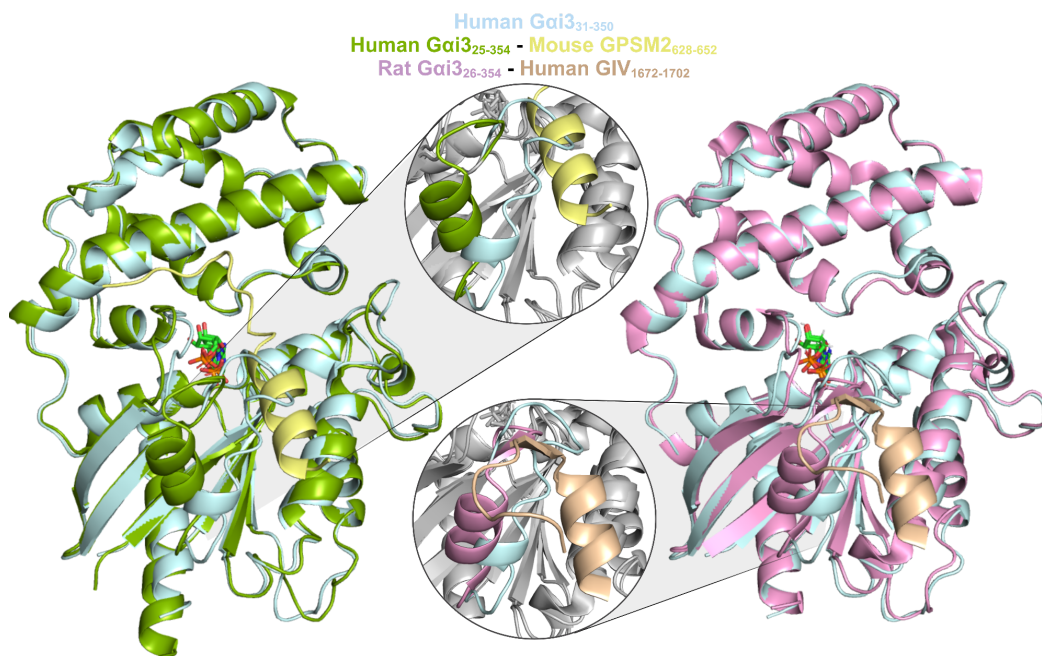
**Figure 33.** Crystal structure of sGai3 in complex with GDP (PDB entry 8OY1), resolved at 3.34 Å. (A) Depicts the structure highlighting the helical domain unique to heterotrimeric G-proteins (pink) and the G-domain or Ras-like domain, featuring six beta strands (purple) and five alpha helices (blue), along with the three switch regions (gold). (B) The same crystal structure with the ribbon colored according to the B-factors of the C $\alpha$  atoms, from blue (lowest) to red (highest).

In the protein data bank, there are several structures of human Gai3 constructs bound to other proteins (identified below by their entry ID):

- 2IHB, 2ODE and 2V4Z are crystal structures of complexes with domains (about 150 residue long) of different human RGS proteins (with GAP activity). Gai3 is bound to GDP,  $Mg^{2+}$ , and a tetrafluoroaluminate ion ( $AlF_4^-$ ), which is used to mimic the transition state for GTP hydrolysis. Bound to this ion, Gai3 is considered to be in its activated form. The RGS domains are RGS10, RGS8, and a triple mutant of RGS2 (Soundararajan et al., 2008; Kimple et al., 2009). They bind to a cleft lined by switch III and the middle region of the connection  $\beta 5-\alpha 4$ , a different region from GIV's GBA motif.
- 4G5O, 4G5R and 4G5S are crystal structures of complexes of GDP-Gai3 with the GoLoco motifs 4 or 3 (about 25 residue long) of the murine G-protein-signaling modulator 2 (GPSM2, also known as LGN, with GDI activity). There is no publication describing them. They bind to the same pocket as the GBA motif of GIV does. In 4G5O, Gai3 has the mutation Q147L (at the helical domain).
- 7E9H is the cryoEM structure of the complex with G $\beta 1\gamma 2$  and the glutamate receptor. The helical domain of Gai3 is not modelled, and GDP is absent (Lin et al., 2021).

Comparing the backbone of GDP-sGai3 with 4G5R (GDP-Gai3 bound to the GPSM2 GoLoco 4 fragment, with a resolution of 3.48 Å), shows a very similar structure ( $RMSD_{C\alpha} = 0.81$  Å), with the major difference in the switch II region (**Figure 34, left**). This region is folded as a one-turn helix and a coil segment in the free protein, but folded into a two-turn helix with a different orientation when bound to the GoLoco peptide. Comparing the backbone of GDP-sGai3 with 6MHF (rat GDP-Gai3 bound to the GBA sequence of GIV, with a resolution of 2.0 Å), also shows a very similar structure ( $RMSD_{C\alpha} = 0.72$  Å), with the major difference in the switch II region (**Figure 34, right**). The switch II helix is longer and differs in its orientation.

These results suggest that the switch region is flexible in GDP-Gai3 and binding of GoLoco or GBA sequences changes its conformation to a more ordered one. The same might happen upon IGGi-11 binding, but as it is smaller than the peptides it might just reduce mobility without inducing a major conformational change in the switch II region.



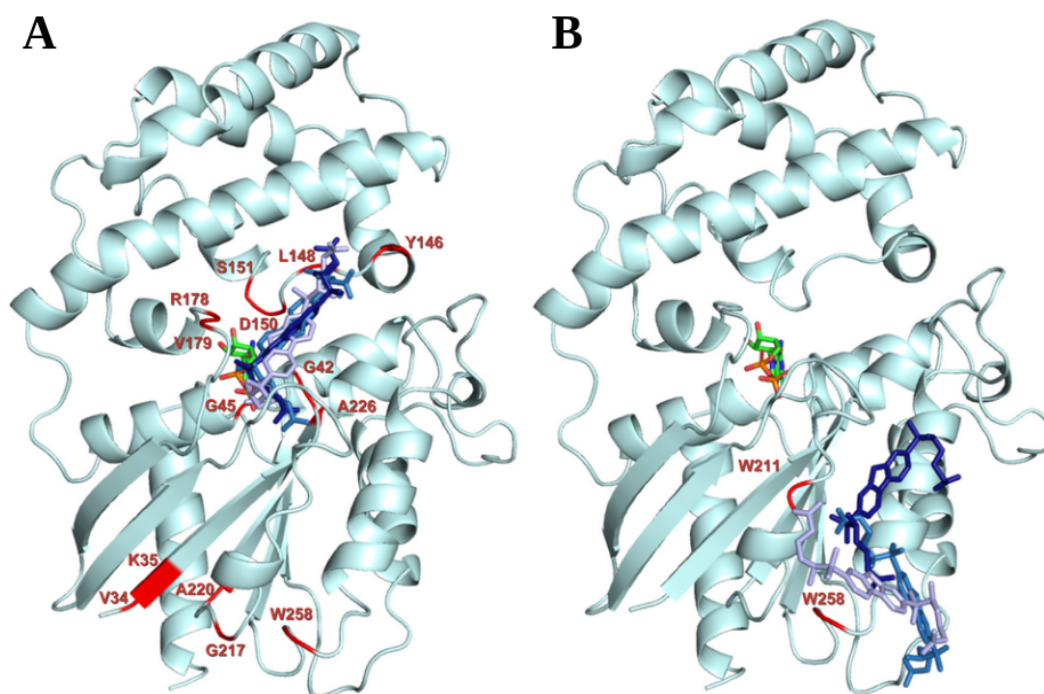
**Figure 34.** *Left panel:* Overlay of the crystal structures of human Gai3 free (PDB entry 8OY1, blue) or bound to the GoLoco 4 motif of GPSM2 (PDB entry 4G5R, green for Gai3 chain and yellow for GPSM2). *Right panel:* Overlay of the crystal structures of human Gai3 (PDB entry 8OY1, blue) and rat Gai3 bound to the GBA motif of GIV (PDB entry 6MHF, pink for Gai3 chain and wheat for GIV). The GDP molecules are shown in sticks.

#### 4.2.5 *In silico* docking of IGGi-11 with sGai3

The crystal structure of sGai3 allows for conducting docking studies with IGGi-11 and gain insight on the intermolecular interactions. This docking might provide complementary information to that from the docking using the homology model built on the structure of human Gai1 bound to a non-natural peptide (Zhao et al., 2023).

##### 4.2.5.1 IGGi-11 docking studies with HADDOCK

Initial attempts used the web server HADDOCK to dock IGGi-11 guided by distance restraints between IGGi-11 and the Gai3 residues experiencing large CSPs. The restraints are ambiguously defined between any atom of IGGi-11 and any atom of those residues (named active residues). This procedure is less costly in computational time, because the force field used to compute the potential energy to rank the different possible docked complexes is simplified (as the experimentally derived restraints are expected to guide the docking and compliance with them is a major contribution to the energy-based ranking). The results, however, were not conclusive.



**Figure 35.** Overlay of IGGi-11 binding poses on sGai3 using HADDOCK with two distinct criteria: (A) based on residues with CSPs exceeding the average plus two times the standard deviation, and (B) based on the tryptophan residues within the G domain. Restrictive residues are highlighted in red, and the top-scoring three poses for each model are displayed in blue.

Using different thresholds to define the active residues (multiples of the median value, as used in the CSP versus sequence plots, or the average plus one or two standard deviations, as used to select residues to calculate the dissociation constant), yielded divergent results but with IGGi-11 in a region near the GDP and interacting with both the helical and the G-domain (**Figure 35, A**). Using as restraints just the two tryptophan residues whose side chain NMR signals are perturbed (W211 and W258, but not 131) resulted in a highly heterogeneous set of IGGi-11 poses, but in the cleft lined by switch II,  $\alpha_3$  and the  $\alpha_3$ - $\beta_5$  loop (where the GIV or GPSM2 peptides bind on rat or human Gai3) (**Figure 35, B**).

We interpret the results as caused by different effects: 1) the CSPs are not only dependent on the intermolecular distance, but also on the orientation of IGGi-11 with respect to the protein, 2) the CSPs may also reflect long distance effects, conformational changes and/or changes in mobility, 3) the energy field of HADDOCK is too simple for a precise docking based only in energy calculations, 4)

a significant portion of switch II and part of  $\alpha 3$  remains unassigned. Assuming IGGi-11 binds similarly to how GIV does, these regions should exhibit the highest CSPs. However, the lack of information in this area lowers the threshold for selecting active residues, leading to the inclusion of residues affected by distance effects as active residues.

#### 4.2.5.2 IGGi-11 docking studies with Autodock

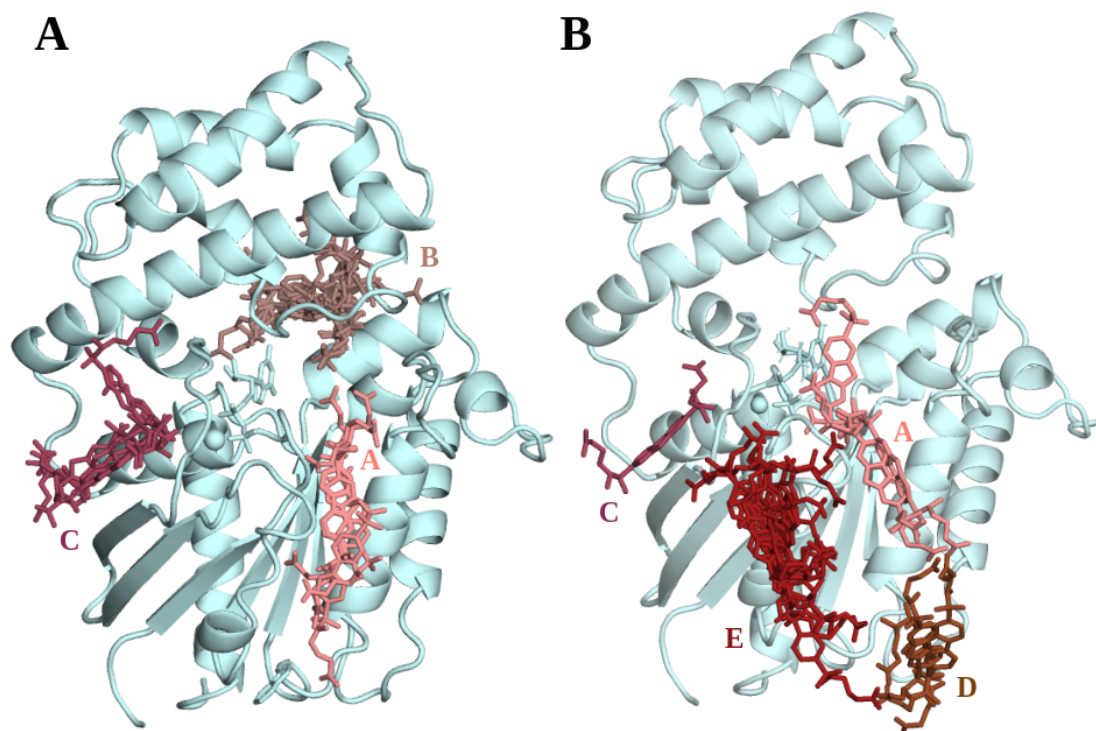
A docking based on a more realistic force field was thus undertaken using AutoDock4 and AutoDock Vina. To avoid exploring all the surface of the protein (which would be computationally very costly), a preliminary search with SiteMap identified two potential binding sites for ligands: one is the cleft lined by switch II, helix  $\alpha 3$  and the N-terminal region of the  $\alpha 3$ - $\beta 5$  loop of sGai3 (named site A), the other is a cleft on the opposite side of the protein, lined by the loop between helices C and D and by helix E of the helical domain, and by the first half of the  $\beta 5$ - $\alpha 4$  region (which is helical and is usually considered as helix F of the helical domain; the corresponding  $\beta 5$ - $\alpha 4$  region in the Ras proteins is a short loop). For docking IGGi-11 on sGai3, the surface inside a square box of 47.25 Å centered around the C $\alpha$  atom of Val225 was explored as it contained the two regions predicted by SiteMap as potential ligand binding sites.

AutoDock Vina docked IGGi-11 to sites A and B, but also to site C, a cleft between the helical and G-domain lined by switch I and the N-terminal region of helix  $\alpha A$  (**Figure 36, A**). Most of the docked molecules were on site B.

AutoDock4 docked IGGi-11 to sites A and C, but also to two sites close to site A. These two other sites do not correspond to cavities, but to the surface of the N-terminal region of  $\alpha 3$ - $\beta 5$  loop and loop  $\alpha 4$ - $\beta 6$  (site D) and to the surface of switch II (site E), mostly formed by polar groups. Most of the IGGi-11 molecules were docked on site E (and a few were docked to other different sites) (**Figure 36, B**).

The possibility of IGGi-11 binding simultaneously to different sites with similar affinity can be discarded based on NMR and ITC data, which indicate that the molar stoichiometry of the complex is 1:1. Yet, the possibility of IGGi-11 binding to one site with measurable affinity (specific binding) and to others with much lower affinity (unspecific binding) cannot be excluded.

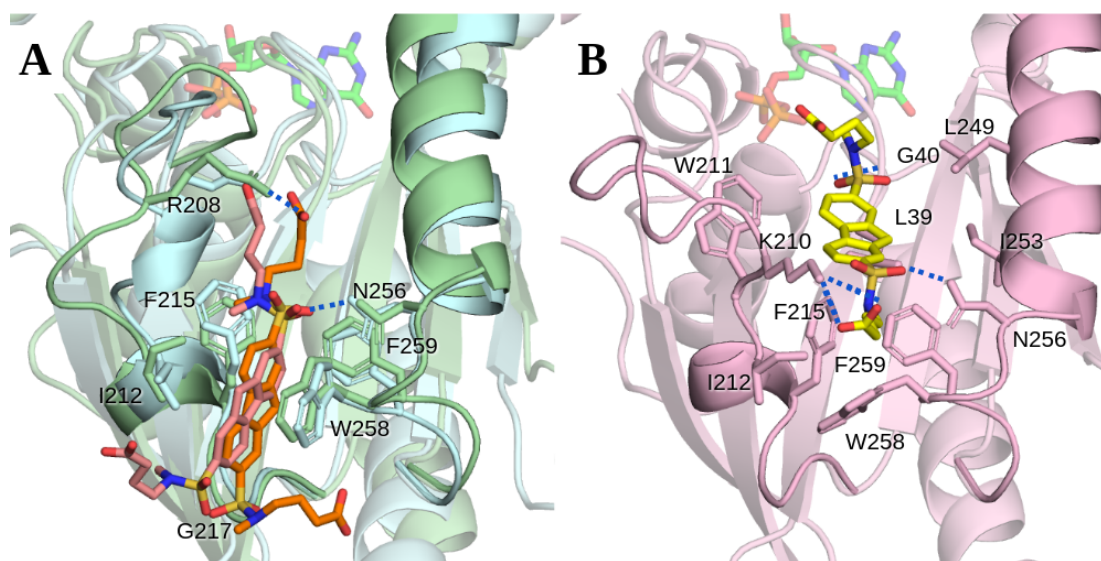




**Figure 36.** Overlay of IGGi-11 binding poses on sGai3. AutoDock Vina (A) yielded 20 binding poses, while AutoDock4 (B) produced 18 binding poses. These conformations can be grouped into five areas labeled A-E, with two of them (A and C) being common to both programs.

Site B is consistent with NMR signal perturbations in the  $\alpha$ CD- $\alpha$ E loop, but it is not with most of the other perturbations (in far-away regions of the G-domain), and is also inconsistent with the absence of perturbations in residues of the helical domain that would be very close to IGGi-11 if it were bound to this site. Site C is not consistent with most of the observed NMR signal perturbations and with absence of many expected perturbations if that were the site of IGGi-11 binding. Site D is consistent with some NMR signal perturbations (W258sc) but inconsistent with most of them (loop  $\alpha$ 1- $\beta$ 1, switch I). Site E would largely perturb the switch II residues, but it is not possible to measure them because most of their NMR signals are not observed or are not assigned. Because the site is superficial and polar it seems unlikely that IGGi-11 binds there, and that would be inconsistent with a 30-fold reduced binding affinity to Gai3 mutant F215A, in switch II but with its hydrophobic side chain at the bottom of the site A cavity. Therefore, sites B, C, D and E are considered unlikely to represent the complex in solution.

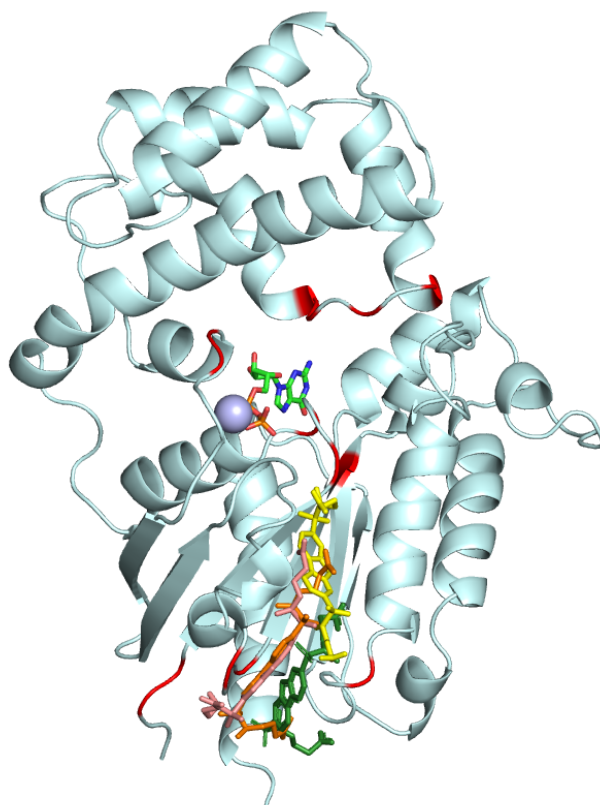
Molecular dynamics simulations were used to further evaluate the probability that IGGi-11 docked to site A best represents the complex. The site A cleft is elongated, and the rigid part of IGGi-11 is also elongated but shorter than the cleft. Although all IGGi-11 molecules docked by Vina in site A are accommodated inside the cleft, some are located with one of the carboxyl groups closer to the switch III loop than others, which in turn have the symmetric carboxyl group closer to the  $\alpha$ 3- $\beta$ 5 loop.



**Figure 37.** Three chosen poses derived from MD simulations of IGGi-11 molecules docked using AutoDock Vina. The interacting residues are highlighted. (A) MD simulation outcomes for docking poses A1.2 (blue) and A2.2 (green), with corresponding MD replicates shown in pink and orange, respectively. (B) MD simulation results for docking pose A2.1 (pink), with the MD replicate represented in yellow.

Two of the poses, one of each kind, were selected for MD simulations: A1, closer to switch III loop, and A2, closer to  $\alpha$ 3- $\beta$ 5 loop, and the dynamics of three identical replicas of each pose were simulated (but with different initial atom velocities). In the six replicas IGGi-11 remained bound to site A along the simulation, with two of the replicas of A1 moving towards the position of the A2 ones. Using the experimental data (mutagenesis and NMR) to select the most likely mode of binding led to the identification of two modes. In both modes IGGi-11 is close to residues F215, N256 and W258 whose mutation decreases binding affinity measured by calorimetry. In mode 1 (represented by replicas A1.2 and A2.2, which converge despite their different initial location on site A) IGGi-11 is closer to the  $\alpha$ 3- $\beta$ 5 loop, with its hydrophobic fluorene moiety of IGGi-11 enclosed between I212 and W258

side chains at either side of the cleft. In mode 2 (represented by replica A2.2) the fluorene contacts residues W211, L249 and I253 (the last two in  $\alpha 3$ ). In mode 1 the carboxylate of the arm whose sulfone interacts with N256) electrostatically interacts with R208, while in mode 2 this carboxylate interacts with K210 at the other end of the elongated cleft (**Figure 37**). Mode 2 is very similar to the docking of IGGi-11 on the homology model of GDP-G $\alpha$ i3 based on the crystal structure of GDP-G $\alpha$ i1 (Zhao et al., 2023), made without molecular dynamics simulations (**Figure 38**).

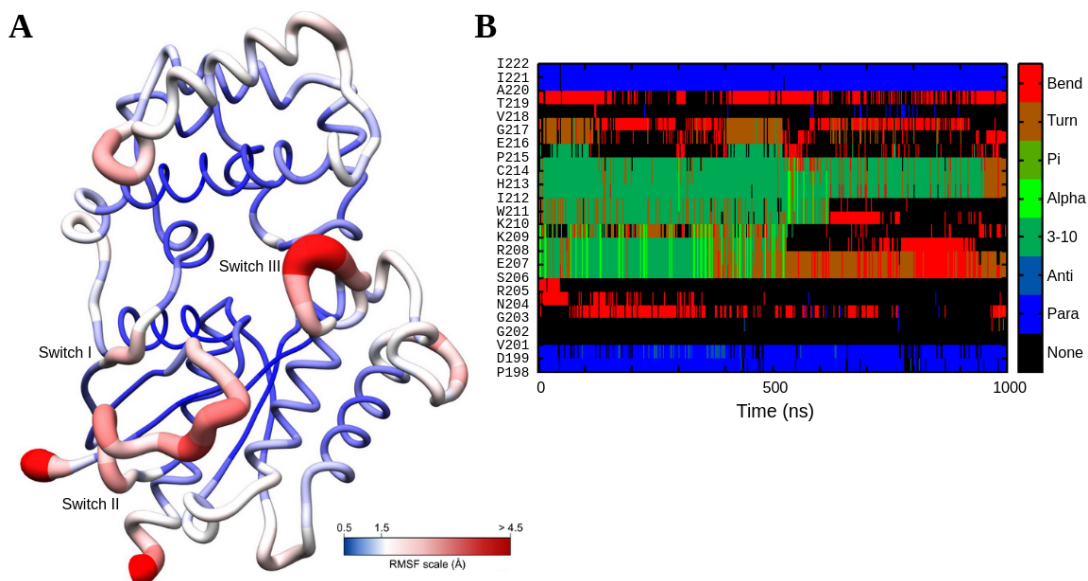


**Figure 38.** The three selected poses from MD simulations of AutoDock Vina (pink, orange and yellow) with the docked IGGi-11 molecule on the homology model of GDP-G $\alpha$ i3 based on GDP-G $\alpha$ i1 (green). Residues with CSPs exceeding the average plus two times the standard deviation are highlighted in red.

We cannot definitively conclude which of the binding modes is prevalent in solution, except that it occurs at site A. However, the length of the cavity, the symmetry of IGGi-11 and the dynamics of switch II suggests that IGGi-11 may bind in different conformations inside the cleft. This would be consistent with NMR signal CSP not confined to a small region in site A, and would not be inconsistent with calorimetry

data as the stoichiometry would still be 1:1, with a measured affinity that would be the weighted-average over the different bound conformations. If this multiple binding mode is correct, then the exchange rate between the different IGGi-11 bound conformations must be very fast (as compared with the chemical shift time scale) because no broadening occurs in the signals of the protein that shift in the complex (at least by visual inspection). Considering a difference of a few tenths of ppm (a few Hz in frequency units) in the <sup>1</sup>H signals of the protein between states with nearby or distant IGGi-11 then the exchange rate must be much faster than the ms time scale.

There are two regions of the protein (besides the chain termini) displaying a high relative dynamic behavior in the MD simulations of the complexes: switch II and switch III. An MD simulation of the protein without IGGi-11 yields the same result: the two switch regions (and the chain termini) are highly dynamic (**Figure 39**). Switch II loses its helical structure after 0.55 μs, but it remains partially helical in the complexes with IGGi-11, suggesting that the ligand reduces its dynamics.



**Figure 39.** (A) Representation of the fluctuations for the C $\alpha$  atoms along the 1  $\mu$ s molecular dynamics trajectory of GDP-Gai3. The thicker the coil the higher the fluctuations. (B) Time evolution of the secondary structure (according to the DSSP classification) of residues in the switch II region.

The heterogeneity of IGGi-11 binding to site A could explain the puzzling result that sGai3 crystals grew in the presence of IGGi-11, but not in its absence, and without co-crystallized IGGi-11. In the crystallization drops, sGai3 is saturated with IGGi-11,

---

and the reduced dynamics may facilitate crystallization, but instead of the crystal sequestering one of the binding modes in solution, it causes dissociation of the heterogeneously bound IGGi-11.

Breaking the symmetry of the IGGi-11 by chemical modification of one of its arms could favor a more specific binding. In the three selected docked poses the interaction between one of the sulfone groups and N256 is maintained, the fluorene is interacting with different hydrophobic groups, and only one of the terminal carboxylates makes electrostatic interactions with basic residues while the other does not. Changing one of the carboxylates by a positively charged group could favor the formation of a cation- $\pi$  interaction with the side chains of W258 or with W211 increasing the affinity of the interaction.

### 4.3 Study the binding of other regulators to Gai3 by NMR

Gβ1γ2 forms a heterotrimer with GDP-Gai3. The structure is not known, but the crystal structure of rat Gai1 bound to Gβ1γ2 has been determined. The affinity of murine Gai1 for Gβ1γ2 is in the nM range.

The fragment 1185-1221 of RGS12 (which contains the GoLoco motif, responsible for its GDI activity) binds GDP-Gai3 with a  $K_D$  of  $54 \pm 4$  nM at room temperature. Mutagenesis experiments suggest that its binding site partially overlaps with that of the GBA motif of GIV (de Opakua et al., 2017).

The non-natural peptide R6A demonstrates a binding affinity to rat GDP-Gai1 with a  $K_D$  of 60 nM at 25 °C and also exhibits GDI activity. The binding site is not known.

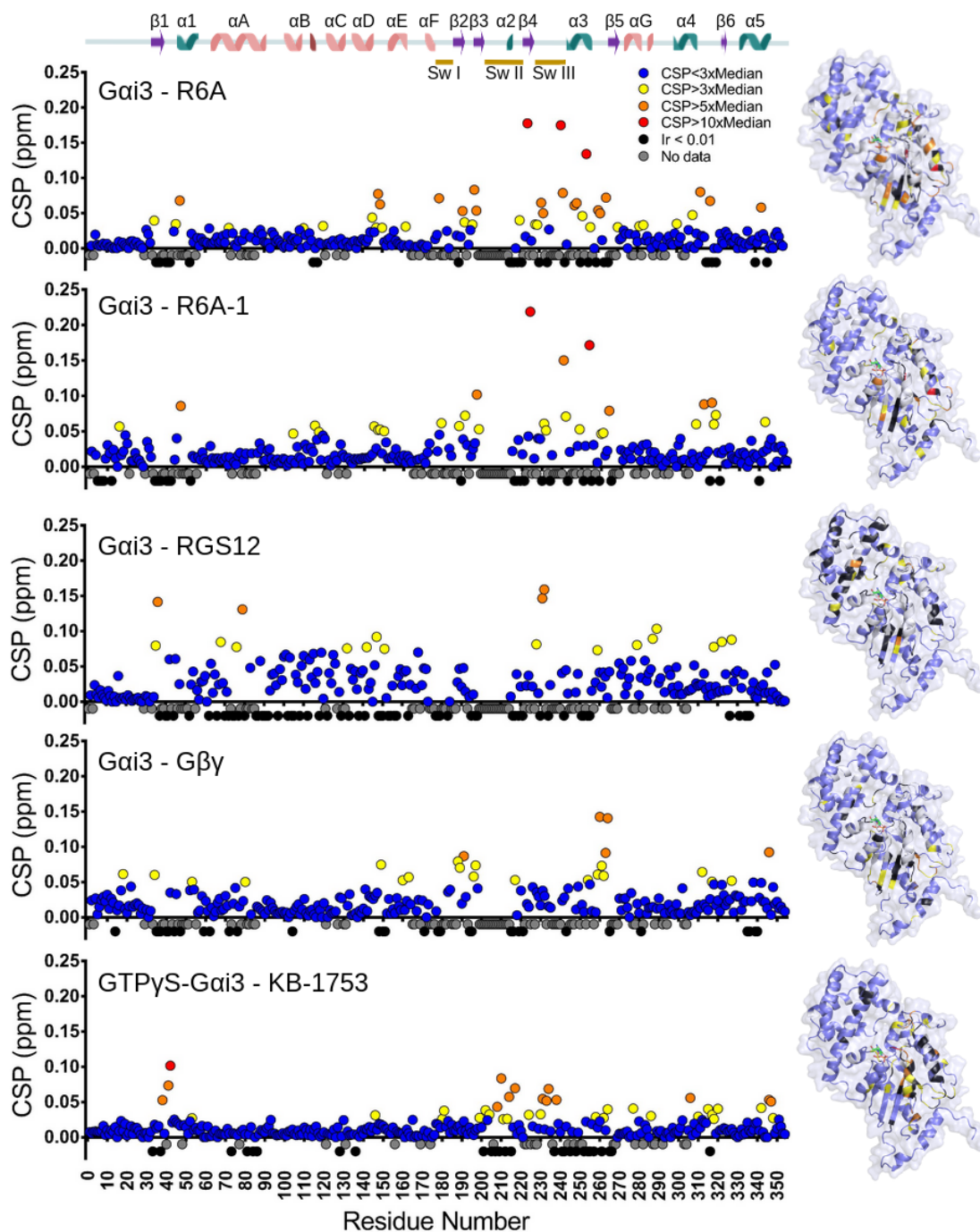
The C-terminal region of R6A, R6A-1, is capable of binding to human GDP-Gai3. Given its affinity for rat GDP-Gai1 (with a  $K_D$  of 200 nM at 25 °C), it is likely that the  $K_D$  for R6A-1 falls within the high nM range. The binding sites are unknown.

The non-natural peptide KB-1753 binds to the switch II-α3 cleft on human GDP-ALF4-Gai1 with a  $K_D$  of 1.2 μM and to human GTPγS-Gai1 with a  $K_D$  of 2.5 μM at 25 °C.

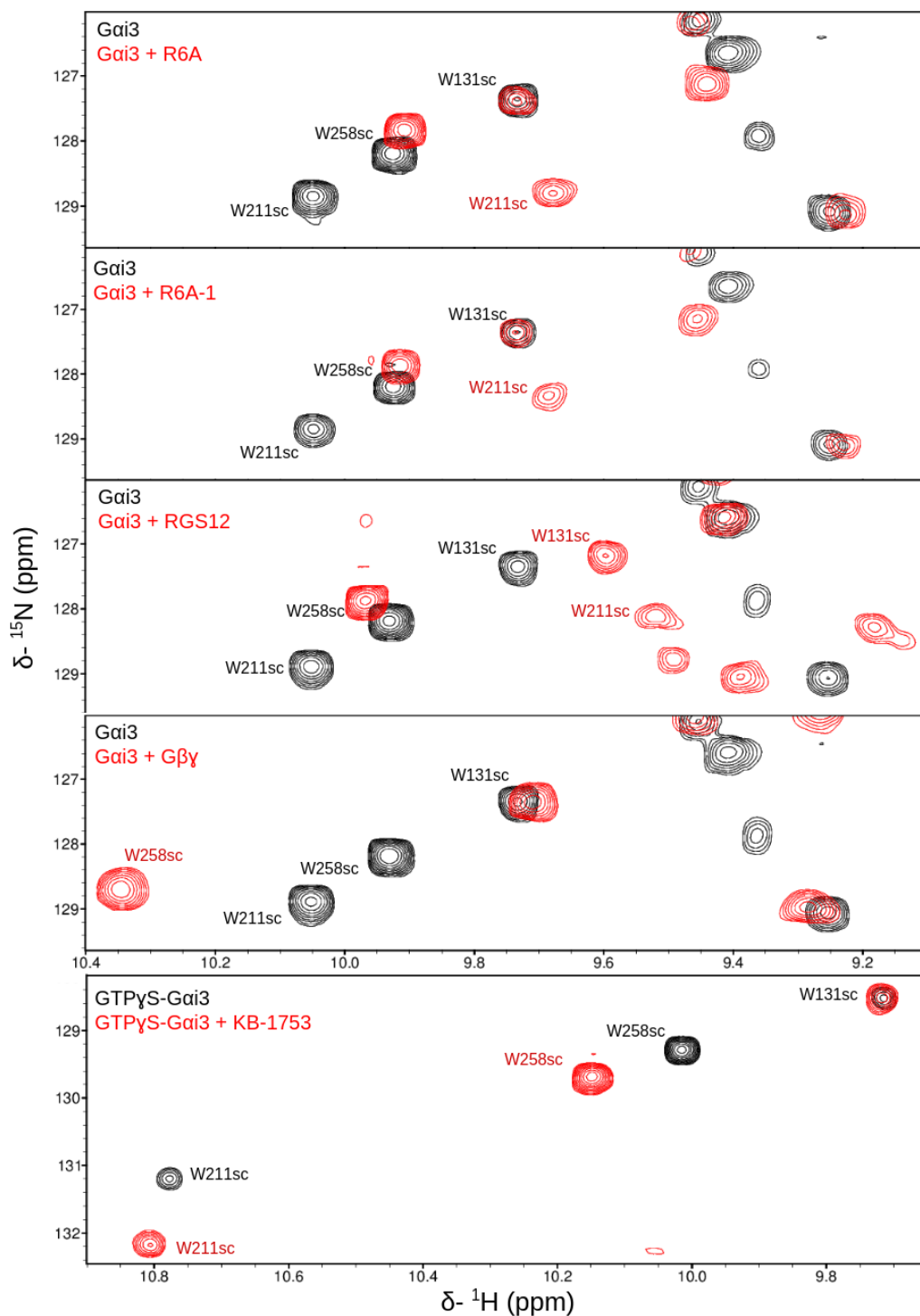
#### 4.3.1 Insights into binding sites of regulators of Gai3

The NMR spectra were recorded under conditions where Gai3 is saturated by all the regulators, according to the estimated  $K_D$  or to the measured CSP along a titration. The analysis of the NMR signal perturbations was done in a similar way as for GIV or IGGi-11: examining the CSP values of the backbone amide (**Figure 40**) as well as the perturbations in the indole signals of three tryptophan residues: W131 in the helical domain, W211 in the switch II region, and W258 in the α3-β5 loop of Gai3 (**Figure 41**). To assign the three indole signals in their bound states, spectra of the complexes with Gai3 or with Gai3 W258F mutant were registered and compared.

The mapping of the backbone amide NMR signal CSPs onto the structural model of Gai3 reveal that all of these regulators bind to the G-domain of Gai3, but RGS12 also binds to the helical domain of Gai3, indicating a larger interaction interface.



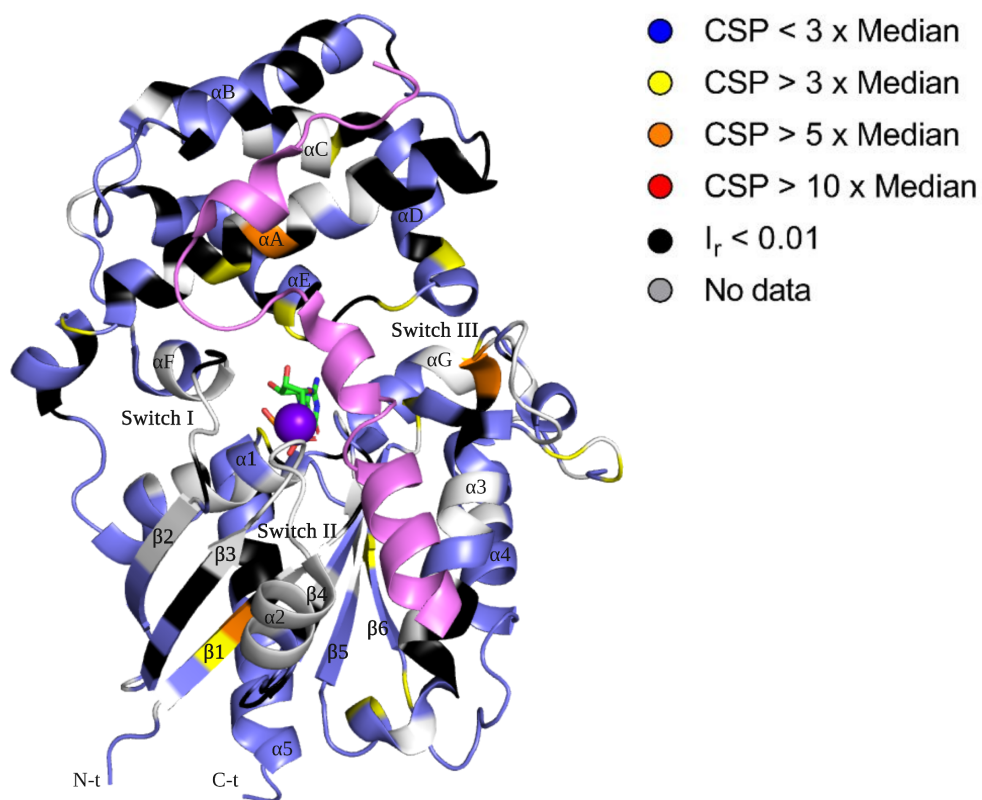
**Figure 40.** Mapping CSPs of *Gai3* interactions with various regulators of *Gai3* (R6A, RGS12, and the  $G\beta 1\gamma 2$  dimer). The CSPs are overlaid on a 3D model of *Gai3*, generated from backbone chemical shifts and with a docked GDP (although the experiments with KB-1753 were done with GTP $\gamma$ S-*Gai3*).



**Figure 41.** Overlay of the region of TROSY spectra of Gai3 containing the tryptophan side chain indole signals (Wsc), alone (black) or in the presence of different regulators (red).



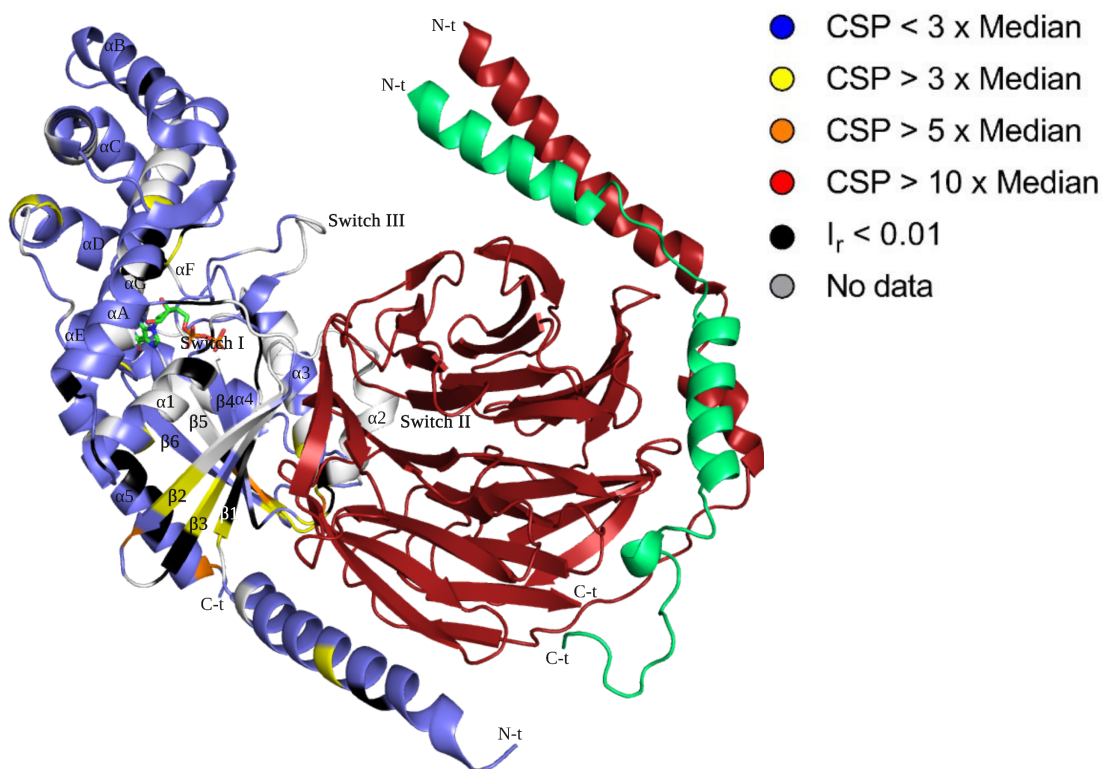
Binding of the 4.4 kDa RGS12<sub>1185-1221</sub> fragment to Gai3 results in partial protein precipitation inside the NMR tube. As a consequence, twelve weak signals become undetectable in the spectrum of the complex, representing 20% of the disappeared signals. The pattern of CSPs show that this fragment containing the GoLoco motif binds to both the G-domain and the helical domain. W211sc (switch II) experiences a very large perturbation while the perturbation on W258sc ( $\alpha 3$ - $\beta 5$  loop) is small. This indicates that RGS12<sub>1185-1221</sub> binds to the same elongated cleft as GIV<sub>1671-1705</sub> peptide but not as close to the  $\alpha 3$ - $\beta 5$  loop (at one the two narrow ends of the cleft). The perturbation on W131sc ( $\alpha C$ ) is also a large one. These results are consistent with the crystal structure of human GDP-Gai1 bound to rat RGS14<sub>496-531</sub> (PDB 1KJY) where the C-terminal half of the peptide binds to cleft reaching the interdomain region where the GDP is bound, and the peptide N-terminal half binds to the helical domain (**Figure 42**). This extensive interaction may also explain why the median CSP caused by RGS12<sub>1185-1221</sub> is larger than those caused by any of the ligands studied.



**Figure 42.** Crystal structure of the human Gai1 bound to rat RGS14<sub>496-531</sub> (PDB 1KJY) is shown. The CSPs observed for the interaction between Gai3 and RGS12 are mapped on the structure and the RGS14 peptide is colored pink.

Binding of the 46 kDa Gβ1γ2 dimer to Gai3 reduced the NMR signal intensity in the spectrum of Gai3, due to the slower rotational diffusion in the bound form and to the large number of protons in non-deuterated Gβ1γ2. Both effects increase the relaxation rate of the Gai3 NMR signals. This loss of sensitivity was partially overcome by increasing the number of scans by a factor of 8 compared with spectra recorded for the small ligands. The pattern of CSP indicates binding to the G-domain. The crystal structure of rat GDP-Gai1β1γ2 (PDB 1GP2) shows the propeller domain of Gβ1 interacting with the switch II region, strands β1 and β3, and the N-terminal region, which is folded as a long helix (**Figure 43**).

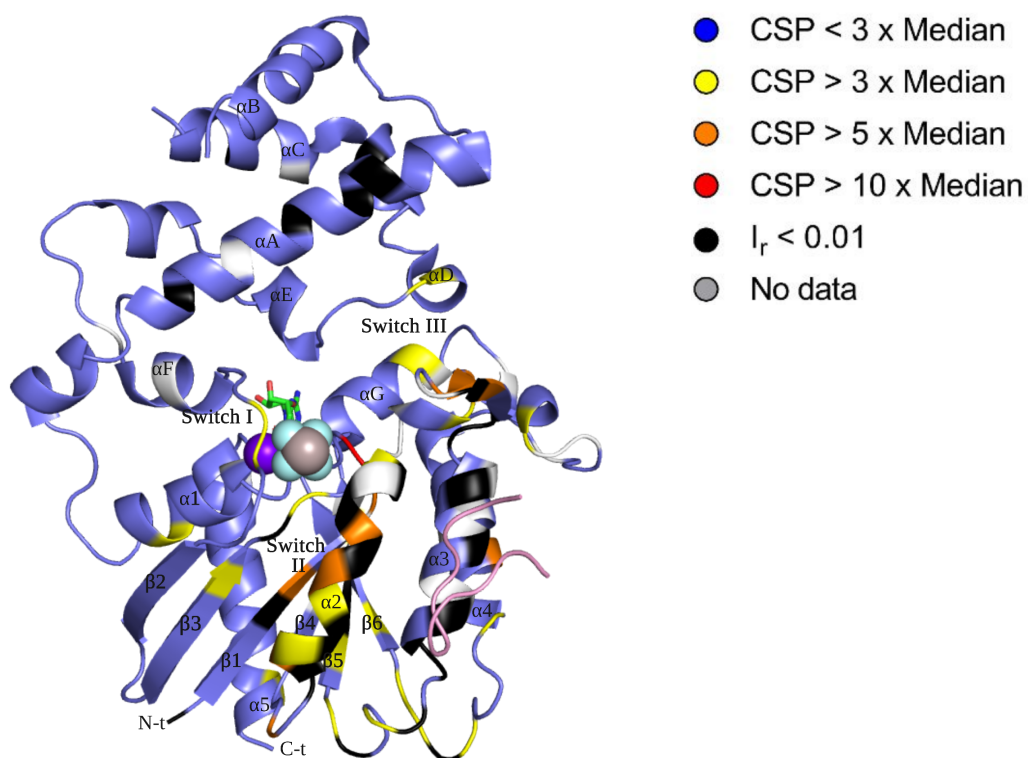
Most of the NMR signals of the switch II region are not detected in the spectrum of GDP-Gai3β1γ2, but several signals in strands β1, β3, and also in β5 (close to switch II) exhibit large perturbations. However, the perturbations in the N-terminal region are small.



**Figure 43.** Crystal structure of the rat GDP-Gai1Gβ1γ2 heterotrimer (PDB 1GP2) is shown. The β-propeller domain of Gβ is depicted in dark red and the γ subunit in green. The CSPs observed for the interaction between GDP-Gai3 and Gβ1γ2 are mapped on the structure of Gai1.

These results indicate that G $\beta$ 1 binds human Gai3 in a similar way as it binds rat G $\alpha$ 1, but with a weak interaction with the N-terminal region. The perturbations on the NMR signals of the side chain of the tryptophan residues agree with this overall picture: W131sc, is slightly perturbed, suggesting that the tight binding of the large G $\beta\gamma$  subunit to the G-domain may induce a long-range, small conformational change in the helical domain; W211sc disappears, in agreement with tight binding of G $\beta$ 1 to the switch II region of Gai3; W258sc, suffers a large perturbation, consistent with being very close to W332 in G $\beta$ 1.

Peptide KB-1753 is specific for the GTP-bound form of Gai subunits. The pattern of CSP measured on GTP $\gamma$ S-Gai3 shows binding to the switch II- $\alpha$ 3 cleft on the G-domain. The signals of W211sc (switch II) and W258sc ( $\alpha$ 3- $\beta$ 5 loop) are similarly perturbed while W131sc ( $\alpha$ C) is unperturbed. These results are consistent with the crystal structure of the complex with GDP-AlF<sub>4</sub>-Gai1 (PDB 2G83), showing the peptide as a  $\beta$ -hairpin with its C-terminal strand inserted into the cleft (**Figure 44**).



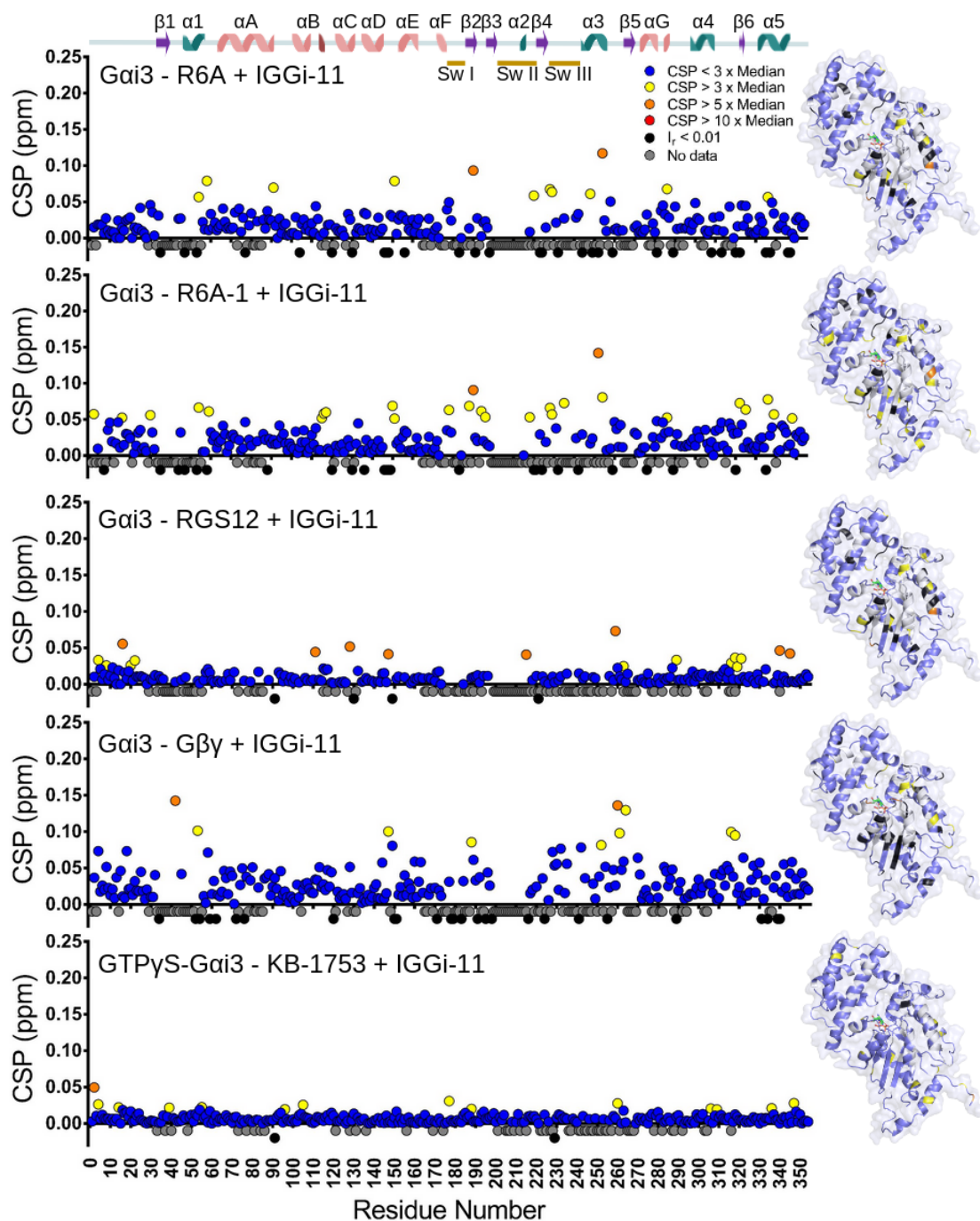
**Figure 44.** Crystal structure of the human GDP-AlF<sub>4</sub>-Gai1 bound to peptide KB-1753 (PDB 2G83). The CSPs observed for the interaction of the peptide with GTP $\gamma$ S-Gai3 are mapped on the structure and KB-153 is colored pink. The Mg<sup>2+</sup> ion is violet and the tetrafluoroaluminate ion is cyan (fluor) and grey (aluminium).

The pattern of CSP of GDP-Gai3 NMR signals is very similar for peptides R6A and R6A-1, both binding to the G-domain and not to the helical domain. W211sc is strongly perturbed, W258sc is slightly perturbed, and W131sc is unperturbed. The results indicate binding to the same cleft on Gai3 where GIV and RGS12 fragments bind, but with the peptide distant from the  $\alpha$ 3- $\beta$ 5 loop, as happens with the RGS12 fragment. This is consistent with competition studies indicating that R6A-1 peptide binds to the same site as peptide KB-752 binds to human GDP-Gai1 (the switch II- $\alpha$ 3 cleft; PDB 1Y3A). The main difference between R6A and R6A-1 peptide binding to GDP-Gai3 observed by NMR occurs at the N-terminal domain. Several residues in this region are perturbed by R6A-1 to a similar extent as by G $\beta$ 1 $\gamma$ 2, while perturbations caused by R6A are very small, similar to those caused by the other ligands, statistically non-significant. On one hand this is surprising because R6A-1 is a 9-residue-long peptide, too short to simultaneously interact with the switch II- $\alpha$ 3 cleft and with the N-terminal region. On the other hand, it has been suggested that the N-terminal 35 residues of human Gai1 are important for both R6A and R6A-1 binding. A possible explanation is that more than one molecule of R6A-1 interacts with Gai3 on different sites and with different affinity. Calorimetry measurements could be used to test this possibility.

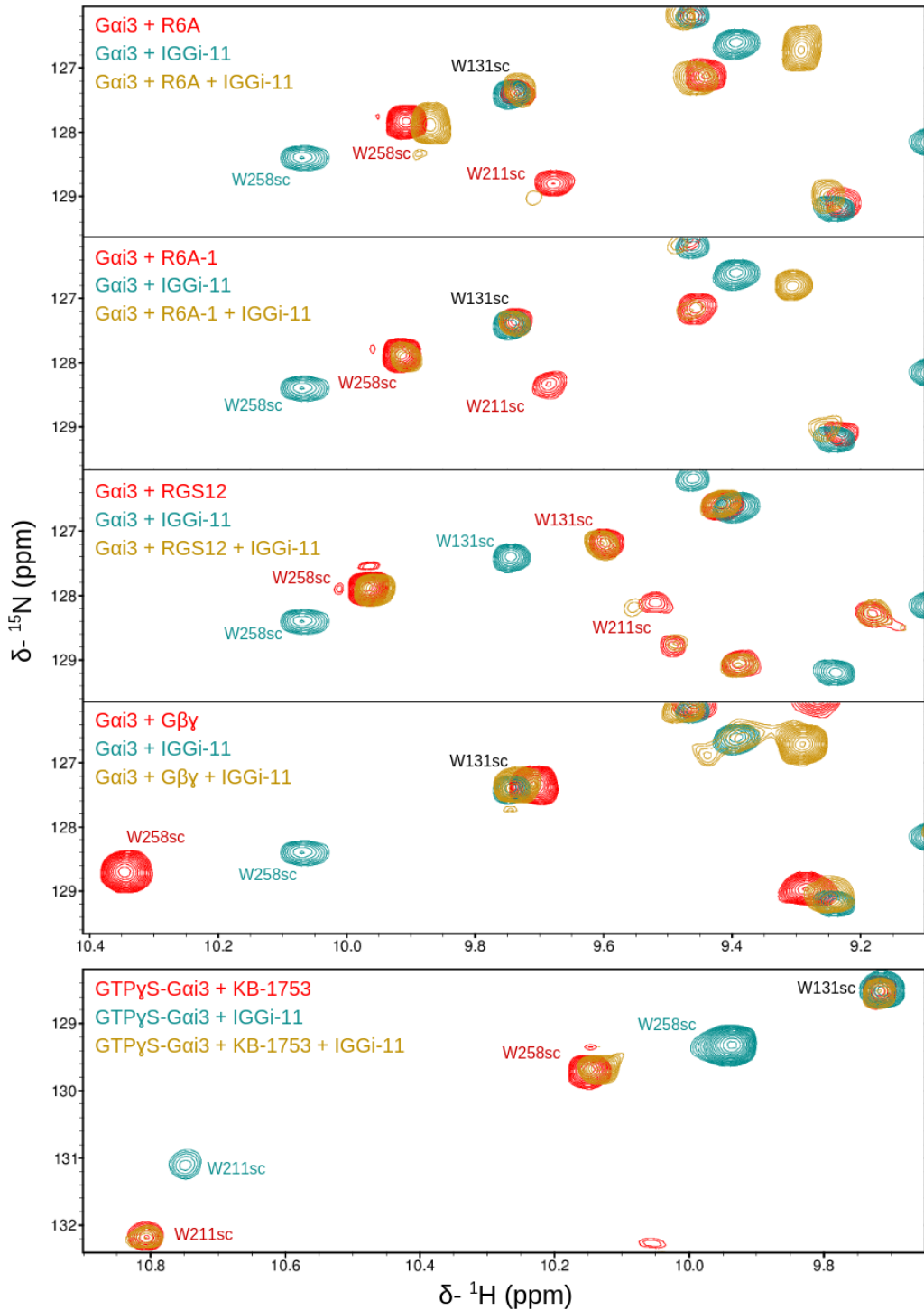
#### 4.3.2 Effect of IGGi-1 on the Gai3 complexes

We aimed to explore the effect of IGGi-1 on the complexes of Gai3 with the studied regulators. To this end, after registering the NMR spectra of the complexes, a few microliters of a concentrated stock of IGGi-1 in DMSO were added directly inside the NMR tube and mixed. The amount of IGGi-1 that can be used in these experiments is limited by the relatively high protein concentration necessary for NMR observation, the limited solubility of IGGi-1 in aqueous buffer, and the destabilizing effect of DMSO on proteins. The [IGGi-1]/[ligand] ratios were: 24 for G $\beta$ 1 $\gamma$ 2, 8 for RGS12, 23 for R6A, 32 for R6A-1, and 10 for KB-1753 (it was 16 for GIV). The final concentration of DMSO was between 1.7 and 3.9 %.

The spectrum of the ternary mixture was compared with the spectra of the binary mixtures to evaluate if IGGi-1 could or could not displace the ligand. The CSP caused on Gai3 by the addition of IGGi-1 to the preformed complexes are plotted in **Figure 45**. And the spectral regions containing the tryptophan residues are shown in **Figure 46**.



**Figure 45.** CSPs of backbone amide NMR signals in the different Gai3 complexes induced by IGGi-11. The CSPs are mapped on the structural model of Gai3.



**Figure 46.** Overlay of the region of TROSY spectra of Gai3 containing the tryptophan side chain indole signals (Wsc) in the presence of different regulators (red), IGGi-11 (blue) and both (gold).

---

Under the conditions used, IGGi-11 does not displace any of the bound ligands. Based on the nM affinity of G $\beta$ 1 $\gamma$ 2, RGS12<sub>1185-1221</sub>, R6A, and R6A-1, ligands for GDP-G $\alpha$ i3, it is not unexpected that IGGi-11 (with a K<sub>D</sub> of 5  $\mu$ M at 30 °C) did not displace them despite binding to the same site (or overlapping sites in the case of RGS12<sub>1185-1221</sub>). And based on the low affinity of IGGi-11 for GTP $\gamma$ S-G $\alpha$ i3, it is expected that it does not displace KB-1753. However, IGGi-11 induces CSPs on GDP-G $\alpha$ i3 bound to G $\beta$ 1 $\gamma$ 2, R6A, and R6A-1, indicating that a ternary complex is formed.

The binding site of IGGi-11 in the ternary complexes is uncertain. However, the CSPs induced by IGGi-11 on G $\alpha$ i3 bound to ligands R6A, R6A-1, and G $\beta$ 1 $\gamma$ 2 indicate that IGGi-11 binds within the G-domain.

It is possible that the hydrophobic IGGi-11 molecule still has some affinity for part of the hydrophobic cleft between switch II and  $\alpha$ 3 in the GDP-bound state but not in the GTP $\gamma$ S-bound one (and thus it does not bind to the complex with KB-1753). It is tempting to speculate that RGS12<sub>1185-1221</sub> binding to GDP-G $\alpha$ i3 changes the conformation and/or dynamics of the bound protein strongly reducing its affinity for IGGi-11.

#### 4.4 NMR-based measurement Gai3's nucleotide exchange rate

The rate of guanosine nucleotide exchange on purified G $\alpha$  proteins has been traditionally measured by a radioligand assay, set up in 1982 by the pioneers of the G $\alpha$  protein biochemistry field using <sup>35</sup>S-GTP $\gamma$ S, which binds G $\alpha$  proteins with same affinity as GTP but it is not hydrolyzed (Sternweis et al., 1981). Using a large excess of GTP $\gamma$ S, the rate of GTP $\gamma$ S binding is limited by the rate of dissociation of GDP from the G $\alpha$  protein. A GDP-G $\alpha$  solution is mixed with a 10-fold molar excess of <sup>35</sup>S-GTP $\gamma$ S, aliquots at different times are filtered and washed (to remove the free nucleotide), and the protein-bound nucleotide (retained in the filter) is quantified by radioactivity. This assay is the gold standard, a direct measurement, and very sensitive (feasible with a few pmols of protein). The disadvantages are that it requires the use of radioactivity and careful sample manipulation. With this assay an exchange rate of  $4.5 \cdot 10^{-4} \text{ s}^{-1}$  at 30 °C has been measured for rat Gai1 (Kleuss et al., 1994).

A related method measures the exchange rate indirectly through the increase in the rate of GTP hydrolysis under the steady-state approximation of enzyme kinetics. This approximation assumes that the concentration of the intermediate state of the catalytic reaction (the GTP-G $\alpha$  complex) does not change with time. This assumption is approximately correct when the concentration of the substrate (GTP) is in a large excess over the enzyme (typically  $[\text{GTP}]/[\text{GDP-Gai3}] = 10$ ). In this condition, the rate of GTP hydrolysis depends only on the rate of exchange of the Gai3-bound GDP for free GTP. The GDP-bound protein is mixed with <sup>32</sup>P-GTP, aliquots at different times are mixed with activated charcoal (to adsorb the protein), centrifuged, and the radioactivity of the supernatant yields the amount of inorganic phosphate produced in the reaction. With this assay an exchange rate of  $2.0 \cdot 10^{-4} \text{ s}^{-1}$  at 30 °C has been measured for human Gai3 (Garcia-Marcos et al., 2009), which increases by a factor of approximately 2.5 when GIV fragments containing the GBA motif are present in a 10-fold molar excess over Gai3 (Garcia-Marcos et al., 2009). This assay shows that IGGi-11 reduces the increase in the exchange caused by GIV (at  $[\text{IGGi-11}]:[\text{GIV}]$  ratios above approximately 10), and that IGGi-11 does not affect the steady-state GTPase activity of Gai3 (below ratios approximately 100), and therefore is not a GDI (Zhao et al., 2023).

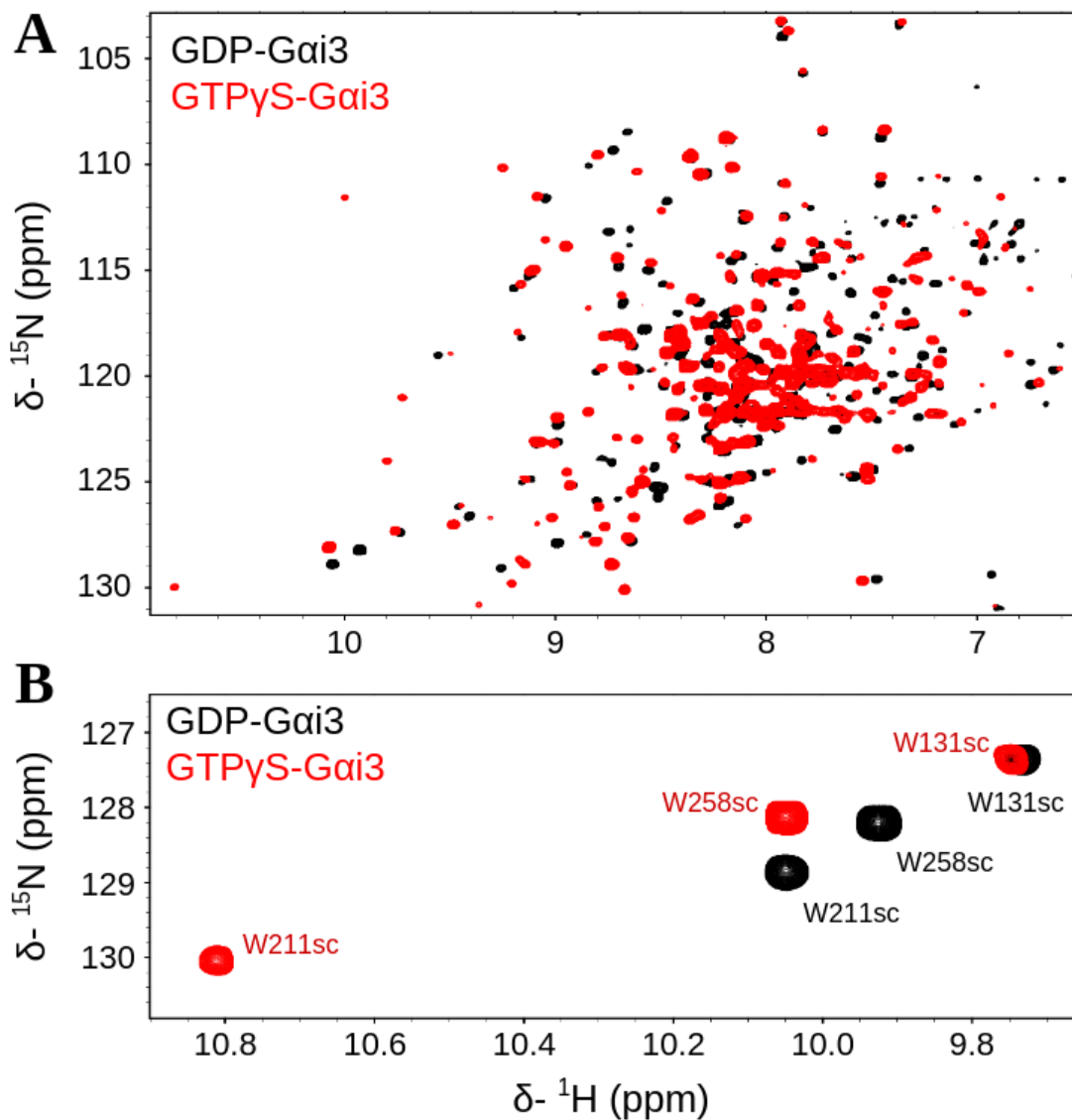
Fluorescent analogs of GTP $\gamma$ S can be used to measure their binding to G $\alpha$  proteins avoiding radioactivity (McEwen et al., 2001). The most popular analog is BODIPY



FL GTP $\gamma$ S, with a bulky fluorophore at the sulfur atom whose fluorescence increases when the modified nucleotide binds to G $\alpha$  proteins. It binds G $\alpha$ i1 protein with approximately 100 times less affinity than GTP $\gamma$ S or GTP, which bind with nM affinity in the presence of Mg<sup>2+</sup>. This is likely the reason the fluorescence assays are not used to measure exchange rates, but to monitor changes caused by regulators of G $\alpha$  proteins. However, the bulky fluorophore can create artifacts, as happened with peptide R6A-1, which binds BODIPY FL GTP $\gamma$ S-G $\alpha$ i proteins quenching the fluorescence of the fluorophore (which led to wrong conclusion that R6A-1 had GDI activity). IGGi-11, at 150-fold molar concentration over G $\alpha$ i3, has little effect on BODIPY FL GTP $\gamma$ S binding to G $\alpha$ i3, indicating that it does not have GDI activity. It has been reported that the nucleotide exchange rate is 10<sup>-4</sup> s<sup>-1</sup> for human G $\alpha$ i3 at 30 °C (Mase et al., 2014) but the method used to derive this value is not mentioned.

NMR spectra show several protein signals with distinct chemical shifts in the GDP- or GTP $\gamma$ S -bound states of G $\alpha$ i3 (**Figure 47, A**). Particularly useful are the signals of the side chains of tryptophan residues 211 and 258, because they appear in a region of the spectrum with no other signals (**Figure 47, B**), except for the signal of W131sc, which changes very little. Therefore, measuring the changes in the NMR signals of W211sc and W258sc over time in a GDP-G $\alpha$ i3 sample with excess GTP $\gamma$ S could directly yield the exchange rate. The difference in frequency for the W211sc NMR signals in the two states is about 600 s<sup>-1</sup> (100 s<sup>-1</sup> for W258sc), much larger than the measured exchange (10<sup>-4</sup> s<sup>-1</sup>). Because exchange is slow in the chemical shift time scale, signals would be observed in both the GDP- or GTP $\gamma$ S-bound states of G $\alpha$ i3, with intensities proportional to their concentrations. This approach has been used on several small GTPases (Mazhab-Jafari et al., 2010; Gebregiworgis et al., 2018). For human KRas, an exchange rate of 7.7·10<sup>-5</sup> s<sup>-1</sup> (presumably at 20 °C in a buffer containing 5 mM MgCl<sub>2</sub>).

Preliminary tests at different temperatures showed that useful spectra could be obtained on a 65  $\mu$ M GTP $\gamma$ S-G $\alpha$ i3 sample at 20 °C in half-hour, registering only the spectral region containing the tryptophan side chain region (8 ppm on the indirect <sup>15</sup>N-dimension). The temperature was chosen to slow down the exchange process (as was done with the small GTPases) to be able to adequately monitor the kinetics with a series of half-hour spectra recorded in an over-night experiment.

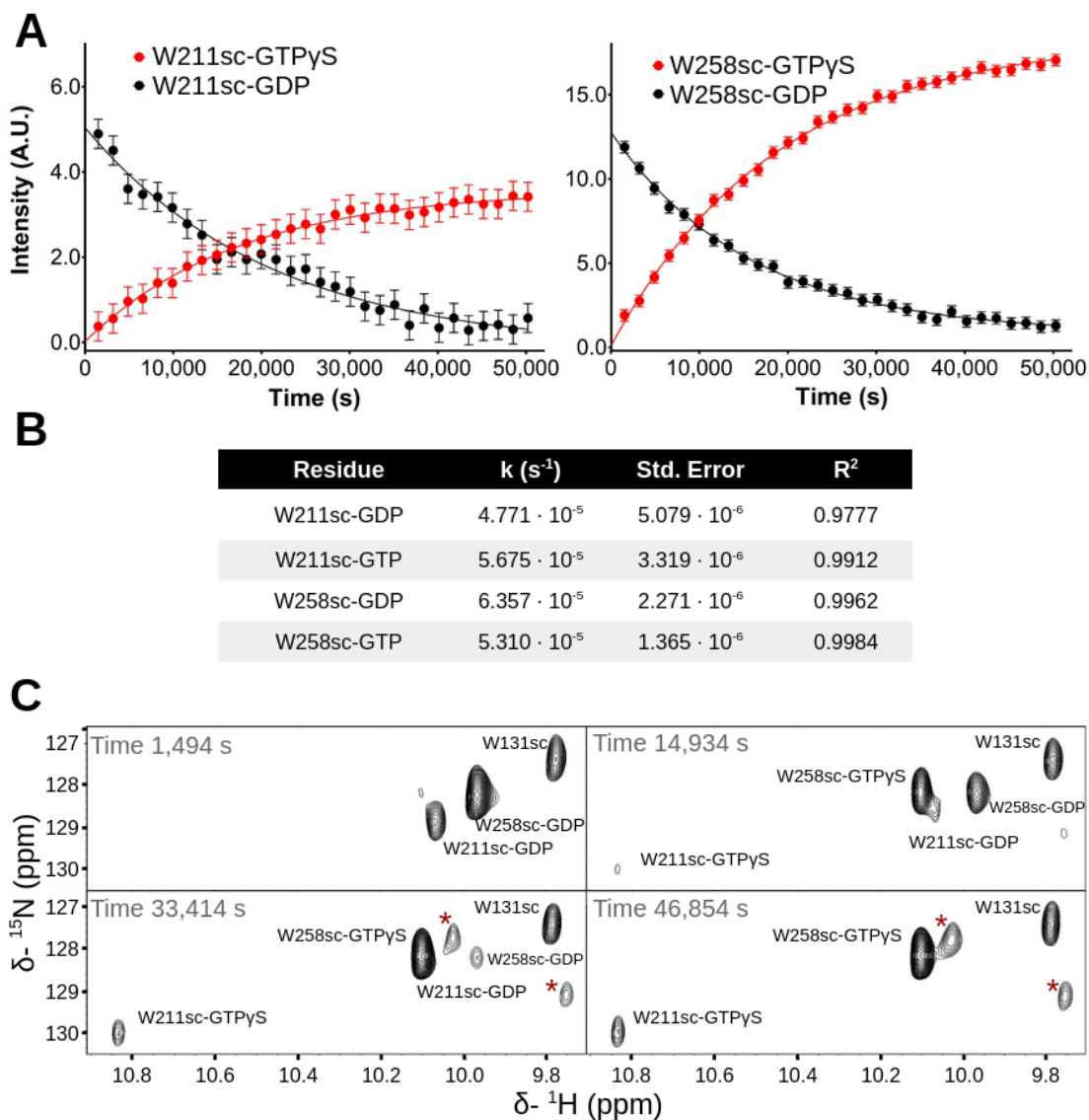


**Figure 47.** (A) Overlay of  $^1\text{H}$ - $^{15}\text{N}$  TROSY NMR spectra of GDP-Gai3 and GTP $\gamma$ S-Gai3 at 30 °C. (B) Region of the spectra with the tryptophan side chain signals.

To apply this NMR-based method, we prepared a 89  $\mu\text{M}$  GDP-Gai3 sample (without excess GDP in the buffer) and then added a 10-fold molar excess of GTP $\gamma$ S. Subsequently, we recorded 30 consecutive  $^1\text{H}$ - $^{15}\text{N}$  TROSY spectra over a total period of 15 hours.

In each spectrum, we measured the intensity of both W211sc and W258sc NMR signals (peak height) for two distinct states: the GDP-bound signals, whose intensity

decreases over time, and the GTP $\gamma$ S-bound signals, whose intensity increases over time, and the data were fit to one exponential decay or to one-phase association, respectively (**Figure 48**).



**Figure 48.** (A) Time-dependent intensity change of W211sc and W258sc signals measured in  $^1\text{H}$ - $^{15}\text{N}$  TROSYTROSSY spectra of GDP-Gai3 after addition of a 10-fold molar excess of GTP $\gamma$ S. The symbols are the experimental measurements and the line is the fitting. The error bars span twice the noise in the spectra. (B) Fitting parameters to equations corresponding to a one exponential decay (GDP-Gai3 signals) or to a one-phase association (GTP $\gamma$ S-Gai3), with  $R^2$  indicating the goodness of fit. (C) Temporal evolution of W211sc and W258sc signals at four distinct time points post GTP $\gamma$ S addition. The red \* represent folded signals, not belonging to this region of the spectrum.

The average of all exchange constants is  $(5.5 \pm 0.7) \cdot 10^{-5} \text{ s}^{-1}$  at 20 °C, which is 0.3 times the rate measured at 30 °C by <sup>32</sup>P radioactivity (García-Marcos et al., 2009). This demonstrates the effectiveness of our method as a direct approach for measuring the nucleotide exchange rate of G*ai*3. In the future, it could be used to assess the impact of different ligands on the exchange rate.

# General Discussion





Human GIV, a large multifunctional protein upregulated in metastatic cancers, interacts with the G $\alpha$ i subunits of trimeric G proteins. Disrupting the GIV-G $\alpha$ i3 interaction with the small molecule IGGi-11 inhibits cell migration without reducing cell viability. It does not interfere with any other major function of G $\alpha$ i3, including guanosine nucleotide binding and hydrolysis, association with G $\beta\gamma$  subunits and other cytoplasmic regulators, activation by GPCRs, or modulation of effectors. These properties make it a lead compound to develop analogs of therapeutic value.

The structural basis of the IGGi-11 inhibition as examined by NMR is that IGGi-11 binds to the same site as GIV fragments containing the GBA motif, demonstrating that it is a competitive (not allosteric) inhibitor. NMR confirms that IGGi-11 is specific for the GDP-bound state of G $\alpha$ i3. This is the inactive state of G $\alpha$ i3, whose slow conversion into the GTP-bound active state is accelerated by GIV. IGGi-11 appears to inhibit specifically signaling cascades mediated by GIV, as occurs with the one initiated by EGFR stimulation.

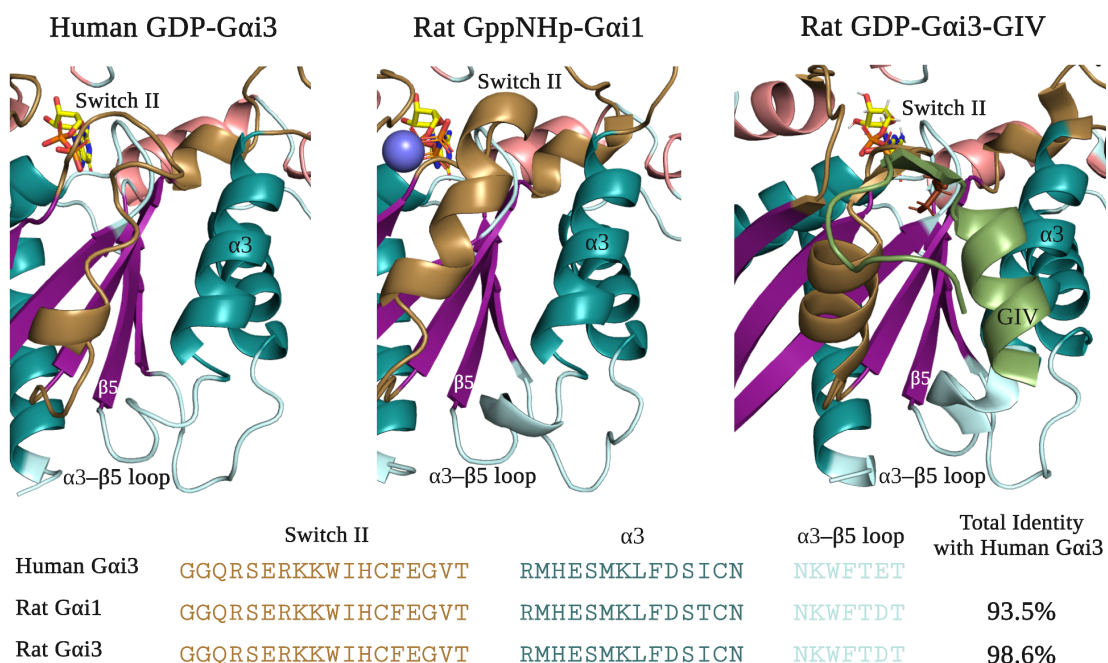
A major limitation of IGGi-11 as an inhibitor is that its affinity for G $\alpha$ i3 is approximately 10 times smaller than the corresponding affinity of GIV. A high-resolution structure of the G $\alpha$ i3-IGGi-11 complex would provide understanding on the molecular recognition and guidance for the synthesis of derivatives with higher affinity. Determination of the structure by NMR is strongly limited by the solubility of the protein. The effort would be very high, and the attainable resolution low, especially in some very relevant regions, since many of the NMR signals in the switch II region are missing in the spectra (likely due to local dynamics unfavorable for NMR signal observation). Precise identification of the most flexible chain termini in solution by NMR allowed the design of a short G $\alpha$ i3 protein, and previous unsuccessful crystallization trials on full-length G $\alpha$ i3 were overcome using this shorter protein. All the chain is visible in the crystal, including the switch II, which is partially helical but with high B-factors. Molecular dynamics simulations confirmed the dynamic nature of this region relative to others.

The crystal structure is highly similar to those of human G $\alpha$ i3 bound to GoLoco motifs of murine proteins (in the switch II-  $\alpha$ 3 cleft) or to the crystal of rat G $\alpha$ i3 bound to GIV. The major difference occurs at switch II, most likely due to the bound ligand. However, co-crystals with IGGi-11 or a GIV peptide were not obtained. This was especially surprising for IGGi-11 since the crystals of short G $\alpha$ i3 were obtained

---

in drops containing saturating concentrations of IGGi-11 (while not crystals were obtained in its absence). It might be that IGGi-11 binding facilitated crystallization but crystals excluded IGGi-11 because they were more stable, or perhaps crystal freezing with cryoprotectant caused IGGi-11 to dissociate from the protein.

Docking IGGi-11 on the crystal structure indicate that its binding to an elongated cleft lined by  $\alpha 3$ , switch II and loop  $\alpha 3$ - $\beta 5$  is heterogeneous, facilitated by the large size of the cleft, and the symmetric nature of IGGi-11. The docked molecules suggest that breaking the symmetry, substituting one of the carboxyl groups by a positively charged group could stabilize one of the possible bound conformers. It is likely, however, that a zero net charge of the molecule would reduce its solubility in aqueous buffers, and alternative polar substitutions should be considered.



**Figure 49:** Zoomed-in view of the region of the cleft lined by the  $\alpha 3$  helix, switch II region, and  $\alpha 3$ - $\beta 5$  loop in human GDP-Gai3, rat Gai1 bound to a non-hydrolysable GTP analog (PDB: 1CIP), and rat GDP-Gai3 bound to the GIV peptide (PDB: 6MHF). The figure illustrates the beta strands (purple), alpha helices (blue), and the switch regions (gold) of the G domain. The magnesium ion is represented as a blue sphere, and the GIV peptide is depicted in green. The sequence of the switch II,  $\alpha 3$  and  $\alpha 3$ - $\beta 5$  loop regions for the three Gai proteins is shown, highlighting their high sequence identity.

The flexibility of switch II in GDP-Gai3 could explain why IGGi-11 (and GIV peptides) are specific for the off-state of Gai3 (they do not bind GTP $\gamma$ S-Gai3, or do



so with an extremely low affinity). In the few G $\alpha$ i proteins crystallized with non-hydrolysable GTP analogs, a large conformational change in the switch II is observed, extending the helix and reducing the size of the cleft. This change most likely restricts the access of ligands that bind there when G $\alpha$ i3 is bound to GDP (**Figure 49**).

Many functional and structural experiments on G $\alpha$ i proteins are done with paralogs or orthologs, and it is assumed that the findings can be extrapolated to a particular G $\alpha$ i interaction. For example, rat homologues are frequently used to evaluate binding to human regulators, by both high (crystallography) or low (fluorescence, and others) resolution techniques.

Studying the binding of human proteins or fragments to human G $\alpha$ i3 by NMR is useful to assess the validity of these measurements and provide additional information. The same occurs with non-natural peptides discovered to bind to G $\alpha$ i proteins. Human GDP-G $\alpha$ i3 binds the GBA motif of GIV in solution in a way consistent with the crystal structure of rat GDP-G $\alpha$ i3 binding to the same motif. Human GDP-G $\alpha$ i3 binds to human G $\beta$ 1 $\gamma$ 2 in solution in a way consistent to rat GDP-G $\alpha$ i1 binding to bovine G $\beta$ 1 $\gamma$ 2 in the crystal structure at 2.3 Å. However, the interaction of G $\beta$ 1 with the N-terminal region of G $\alpha$ i3 (which becomes ordered into a long helix) appears to be weak in solution. Human GDP-G $\alpha$ i3 binds in solution to a fragment of human RGS12 containing the GoLoco motif in a way consistent with the crystal structure of human GDP-G $\alpha$ i1 bound to the homologous fragment of rat RGS14, interacting with both the G-domain and the helical domain.

The non-natural peptides R6A (with GDI activity) and R6A-1 bind G $\alpha$ i proteins, and they bind to human GDP-G $\alpha$ i3 on the same site as the GBA of GIV does, with R6A-1 perhaps also binding to the N-terminal disordered fragment. Because this peptide is very short it is unlikely that a single molecule binds simultaneously to both sites, and further measurements by calorimetry would be necessary to confirm the stoichiometry of the interaction. The non-natural peptide KB-1753 also binds to the cleft lined by  $\alpha$ 3, switch II and loop  $\alpha$ 3- $\beta$ 5 on human GTP $\gamma$ S-G $\alpha$ i3 in solution, as it binds human GDP-AIF<sub>4</sub>-G $\alpha$ i1 in the crystal.

Solution NMR on these complexes after addition of a large excess of IGGi-11 showed that it did not displace any of the bound ligands, as expected based on the relative affinities and concentrations used in these experiments and on previous observations

by other techniques. But perturbations in the NMR signals of some Gai3 residues indicated that IGGi-11 binds to G domain forming a ternary complex. This binding was not observed to GDP-Gai3-RGS12 or to GTP $\gamma$ S-Gai3-KB-1753 complexes.

Previous experiments demonstrating displacement of GIV from Gai3 by IGGi-11 (in pull-down and fluorescence polarization assays) used very high relative concentrations of IGGi-11 ([IGGi-11]/[GIV] = 294 or 1200, respectively), and Gai3 was preincubated with IGGi-11 before adding the GIV peptide fused to GST or the fluorescently labeled GIV peptide. The high relative concentration of the inhibitor, rather than the order of ligand addition, is the likely reason for not observing GIV peptide displacement by IGGi-11 in the NMR experiments (where ([IGGi-11]/[GIV] = 16). Higher relative concentrations are not attainable in NMR experiments due to the limited solubility of IGGi-11 in aqueous buffer.

NMR analysis of the full <sup>1</sup>H-<sup>15</sup>N TROSY spectrum of Gai3 provides a lot of information but is very time consuming, and still many assignment ambiguities cannot be easily solved. A simpler way of obtaining rapid information on ligand binding, albeit with lower resolution, is to exploit the existence of three tryptophan residues that i) are conveniently located in the protein: one in the helical domain (in helix  $\alpha$ C), and two in the G-domain (switch II, and in the  $\alpha$ 3- $\beta$ 5 loop, where many ligands bind, and ii) appearing in a spectral region with few other signals, and ii) their side chain signals appear in a spectral region with very few other signals. This strategy could provide useful and quick information for other ligand binding events.

The tryptophan NMR signal proved very useful to measure the intrinsic guanosine nucleotide exchange rate of human Gai3. The advantage of this method over the gold standard in the field of G proteins is that it is a direct measurement and it does not require radioactivity. The disadvantages are that it requires more amount of protein (nmols instead of pmols), and is isotopically labeled. A further limitation is that sensitivity makes these rate measurements more efficient at temperatures below the physiological one, but this might be of lesser importance when used to evaluate the impact of different regulators.

# Conclusions





---

The conclusions that can be drawn from this work are:

- IGGi-11 binds to the same site on the G domain of human G $\alpha$ i3 as GIV fragments with the GBA motif do, indicating that IGGi-11 is a competitive inhibitor.
- The affinity of IGGi-11 for GTP-G $\alpha$ i3 is 188-fold lower than for GDP-G $\alpha$ i3, therefore IGGi-11 is highly specific for the off-state of G $\alpha$ i3.
- IGGi-11 effectively disrupts the downstream signaling events mediated by the G $\alpha$ i3-GIV interaction in EGF-stimulated MDA-MB-231 cells.
- The N-terminal and C-terminal regions of GDP-G $\alpha$ i3 are highly flexible and remain so when bound to GIV.
- Deletion of the most flexible residues at the G $\alpha$ i3 chain ends does not perturb binding to GIV or IGGi-11.
- The shortened GDP-G $\alpha$ i3 crystalizes with a structure similar to G $\alpha$ i3 structures bound to different regulators, but it does not co-crystalize with GIV or IGGi-11.
- Docking of IGGi-11 to the crystal structure of the shortened GDP-G $\alpha$ i3 reveals details of binding to the cleft lined by  $\alpha$ 3, switch II, and the  $\alpha$ 3- $\beta$ 5 loop of G $\alpha$ i3.
- G $\beta$ 1 $\gamma$ 2 binds to the G domain of GDP-G $\alpha$ i3, but it does not strongly interact with the N-terminal region,
- The binding site of the RGS fragment 1185-1221 containing the GoLoco motif spans the G domain (in the same cleft as fragments of GIV with the GBA motif) and the helical domain of GDP-G $\alpha$ i3.
- The non-natural peptides R6A and R6A-1 bind to GDP-G $\alpha$ i3 on the same site as GIV fragments containing the GBA motif, with R6A-1 possibly interacting also with the N-terminal region.
- The non-natural peptide KB-1753 binds to GTP $\gamma$ S-G $\alpha$ i3 on the same site as GIV fragments containing the GBA motif.
- IGGi-11 is able to bind to G $\alpha$ i3 complexes with G $\beta$ 1 $\gamma$ 2, GIV<sub>1671-1705</sub>, R6A, and R6A-1, but not to those with RGS12<sub>1185-1221</sub> or KB-1753.
- The exchange rate of GDP for GTP $\gamma$ S of G $\alpha$ i3 is  $(5.5 \pm 0.7) \cdot 10^{-5} \text{ s}^{-1}$  at 20 °C, directly measured by a NMR method.





# Spanish Summary





Las proteínas G desempeñan un papel fundamental en la señalización celular al alternar entre dos estados: uno inactivo unido a GDP y otro activo unido a GTP. En su estado activo, estas proteínas regulan una amplia variedad de procesos biológicos. La transición hacia el estado activo implica el intercambio de nucleótidos, mientras que el regreso al estado inactivo se logra mediante la hidrólisis del GTP. A pesar de que las proteínas G tienen la capacidad de llevar a cabo este ciclo por sí solas, suelen contar con una serie de reguladores debido a que las cinética de intercambio y de hidrólisis son lentas. Entre ellos se encuentran los GEF, que aceleran el intercambio de nucleótidos; los GAP, que aceleran la hidrólisis de GTP; y los GDI, que ralentizan la disociación del GDP. Además, según su estructura oligomérica, las proteínas G se dividen en dos categorías: monoméricas y heterotriméricas.

Las proteínas G heterotriméricas se componen de tres subunidades:  $\alpha$ ,  $\beta$  y  $\gamma$ , con la subunidad  $\alpha$  siendo la que se asocia con el nucleótido. Cuando la subunidad  $\alpha$  cambia de GDP a GTP, se disocia del complejo  $\beta\gamma$ , y ambas partes inician sus respectivas cascadas de señalización. Los GEF canónicos para las proteínas G heterotriméricas son los receptores acoplados a proteínas G (GPCRs), que aceleran el intercambio de nucleótidos de guanina en la subunidad  $\alpha$ . Así, las proteínas G heterotriméricas son componentes clave en la transducción de señales desde el exterior de la célula a su interior, a través de la membrana citoplasmática. Sin embargo, algunas de estas proteínas G también pueden ser activadas por proteínas citosólicas.

Una de estas proteínas GEF no receptora es la proteína GIV, que tiene la capacidad de activar  $G\alpha i3$ . Esto desencadena una señalización intracelular que influye en la migración celular, un proceso esencial para la cicatrización de heridas, la respuesta inmunológica y el desarrollo embrionario.

La interacción entre GIV y  $G\alpha i3$  ha sido asociada con cánceres metastásicos. La elevada concentración de GIV en estos tipos de cáncer puede conducir a una regulación desequilibrada de la migración celular y la invasión de tejidos, lo que facilita la diseminación de células cancerosas a otras partes del cuerpo. Por lo tanto, se ha identificado la interrupción de la interacción entre GIV y  $G\alpha i3$  como un posible diana terapéutico para el tratamiento de ciertos tipos de cáncer.

La interacción entre estas moléculas se ha analizado previamente mediante resonancia magnética nuclear (RMN), y se ha demostrado su susceptibilidad a la inhibición farmacológica a través de dos estrategias de detección complementarias. En conjunto,

se lograron identificar ocho compuestos únicos a partir de ambos enfoques. Se realizaron curvas de inhibición en función de la dosis para los ocho compuestos utilizando FP, y se encontró que NF023 presentaba una inhibición dependiente de la dosis con valores de IC<sub>50</sub> cercanos a 5 μM. No obstante, no resulta ser un candidato adecuado para el desarrollo de fármacos, ya que no atraviesa la membrana celular y carece de especificidad. A pesar de esto, estos hallazgos respaldan la posibilidad de dirigirse específicamente a la interacción Gai3-GIV.

Tras este avance, nuestros colaboradores llevaron a cabo un cribado de 200,000 compuestos con el objetivo de identificar un posible inhibidor. En este proceso de cribado, se consideraron diversos factores, como la afinidad de unión a Gai3, la disponibilidad para su adquisición y la viabilidad celular. Entre todos los compuestos analizados, únicamente uno, IGGi-11, demostró tener actividad inhibitoria contra la interacción GIV-Gai3.

En el marco de esta investigación, se llevó a cabo un análisis estructural exhaustivo de la interacción entre IGGi-11 y Gai3, así como de la interacción entre el péptido GIV<sub>1671-1705</sub>, que contiene el motivo GBA de GIV, y Gai3. Dado que Gai3 tiene un peso molecular de 41 kDa, se enriqueció la muestra con deuterio y <sup>15</sup>N de GDP-Gai3. Se realizó una titulación de esta muestra con IGGi-11 y se comparó con el espectro de la interacción entre Gai3 y GIV<sub>1671-1705</sub>.

Los resultados de esta titulación revelaron cambios en los desplazamientos químicos en Gai3 en respuesta a la presencia de IGGi-11, y estos cambios fueron similares a los observados con GIV<sub>1671-1705</sub>. Estos hallazgos sugieren que IGGi-11 actúa como un inhibidor competitivo (no alostérico) en la interacción entre Gai3 y GIV<sub>1671-1705</sub>.

El análisis de la titulación proporcionó una estimación de la constante de disociación, que fue de aproximadamente 6 μM para la interacción Gai3-IGGi-11, una cifra respaldada por resultados similares obtenidos mediante calorimetría.

Dado que GIV se une específicamente a la forma inactiva Gai3, se llevó a cabo una titulación por RMN, similar a la mencionada anteriormente, con el propósito de investigar la interacción entre IGGi-11 y GTP-Gai3. En este caso, las perturbaciones en las señales de RMN causadas por la interacción de IGGi-11 con GTP-Gai3 fueron pequeñas y afectaron las mismas regiones que en el caso de GDP-Gai3. Esto indica que IGGi-11 se une al mismo sitio de unión, pero con una afinidad significativamente

más baja. La afinidad se estimó en  $1014 \pm 235 \mu\text{M}$  a  $30^\circ\text{C}$ , lo que representa una disminución de 188 veces en comparación con la afinidad observada en la interacción con GDP-G $\alpha$ i3. Por lo tanto, IGGi-11 es también específico a la forma inactiva de G $\alpha$ i3.

Se investigó cómo IGGi-11 afecta a las vías de señalización celular reguladas por la interacción G $\alpha$ i3-GIV. Brevemente, tras el estímulo con EGF, GIV se enlaza a EGFR formando el complejo GIV-EGFR junto a G $\alpha$ i3. En células con GIV funcional, esto activa vías de migración y proliferación. Entre ellas, se activa la vía PI3K/Akt, lo que conlleva la activación de proteínas como Akt y sus sustratos, por ejemplo PRAS40. Para analizar el impacto de IGGi-11 en esta cascada de señalización, se emplearon células MDA-MB-231, derivadas del cáncer de mama, que tienen una alta expresión de GIV. Tras el tratamiento con IGGi-11, se observó una reducción en la fosforilación de Akt y PRAS40, lo que sugiere que la cascada de señalización G $\alpha$ i3-GIV se inhibe en presencia del compuesto. Es importante destacar que este efecto no afectó la cantidad total de proteína, lo que indica la especificidad de la acción de IGGi-11 en esta interacción.

La obtención de una estructura cristalina de G $\alpha$ i3 en complejo con IGGi-11 o un péptido GIV sería fundamental para mejorar nuestra comprensión de sus interacciones moleculares. Estas estructuras proporcionarían detalles esenciales sobre el modo de unión y los eventos clave entre G $\alpha$ i3 y el inhibidor o péptido. Esta información permitiría diseñar y optimizar el inhibidor de manera racional, ajustando su afinidad y especificidad. Sin embargo, no ha sido posible cristalizar G $\alpha$ i3 en forma libre o en complejo con ligandos pequeños.

Nuestra hipótesis se centró en que dicha dificultad se debía a la presencia de regiones flexibles que obstaculizaban la cristalización. Por lo tanto, recurrimos a ensayos de relajación mediante RMN de  $^{15}\text{N}$  con el fin de identificar estas áreas flexibles y diseñar una versión más propicia para la cristalización de G $\alpha$ i3 humana. El análisis mediante RMN determinó la existencia de 31 residuos en el extremo N-terminal y 6 residuos en el extremo C-terminal que presentaban alta flexibilidad. En consecuencia, diseñamos la construcción sG $\alpha$ i3, eliminando estas zonas flexibles (G $\alpha$ i3<sub>31-350</sub>) e incorporando una fusión con ubiquitina con una cola de histidinas, lo que facilitó su purificación, aumentó la expresión y mejoró la solubilidad. La ubiquitina se separaba de la proteína mediante un sitio de corte para la proteasa TEV.

Por SEC-MALS se confirmó que sGai3 se encontraba en su forma monomérica con un peso molecular de 37.5 kDa. Mediante ITC a 25 °C, se evaluó la afinidad de IGGi-11 y GIV<sub>1672-1693-DY</sub> por sGai3 y se observó que su afinidad era comparable a la que se encontraba con Gai3 de longitud completa. Tras realizar estos ensayos para asegurarnos de que la construcción era adecuada para los experimentos de cristalización, procedimos con los ensayos de cristalización, tanto con sGai3 en su forma libre como en complejo con GIV<sub>1672-1693-DY</sub> e IGGi-11.

Los ensayos en los que la proteína se encontraba en presencia de un ligando tuvieron éxito en la cristalización, pero no logramos obtener cristales con sGai3 en su forma libre. Tras múltiples etapas de optimización, obtuvimos una densidad electrónica de baja resolución a partir de los cristales de los ensayos en los que sGai3 estaba en presencia de IGGi-11. Al analizar los datos de difracción de estos cristales, nos sorprendió descubrir que proporcionaron la estructura de sGai3 en su forma libre con una resolución de 3.34 Å.

Ante las dificultades en la cristalización, optamos por un enfoque de modelado molecular utilizando HADDOCK. Inicialmente, aplicamos restricciones de distancia basadas en CSPs de la interacción entre IGGi-11 y Gai3. Sin embargo, esta metodología generó resultados ambiguos. Ante la falta de conclusiones claras, se llevó a cabo un docking basado en energías sin restricciones.

Utilizamos SiteMap para identificar dos posibles sitios de unión para los ligandos: uno en la hendidura formada por switch II, la hélice  $\alpha 3$  y la región N-terminal del bucle  $\alpha 3$ - $\beta 5$  de sGai3, y otro en una hendidura en el lado opuesto de la proteína, formada por el bucle entre las hélices C y D, la hélice E del dominio helicoidal y la primera mitad de la región  $\beta 5$ - $\alpha 4$  (a menudo denominada hélice F del dominio helicoidal). Luego, mediante AutoDock, identificamos cinco posibles sitios de unión, pero excluimos cuatro de ellos debido a inconsistencias con datos experimentales de RMN y mutagénesis. Por lo tanto, nos enfocamos en el sitio que involucra switch II, la hélice  $\alpha 3$  y la región N-terminal del bucle  $\alpha 3$ - $\beta 5$  de Gai3, donde encontramos tres poses estables.

Dos de estas poses se agrupan en lo que denominamos "modo 1", que se sitúa más cerca del bucle  $\alpha 3$ - $\beta 5$ , mientras que la tercera se ubica en el "modo 2", donde el núcleo hidrofóbico de fluoreno de IGGi-11 queda más próximo a W211. En ambos modos, el núcleo hidrofóbico de fluoreno de IGGi-11 se encuentra encapsulado en un

bolsillo hidrofóbico de *Gai3* y los grupos carboxilo interactúan con residuos cargados positivamente (R208 en el modo 1 y K210 en el modo 2). Las moléculas acopladas sugieren que modificar la simetría podría estabilizar una de las posibles conformaciones de unión.

Además, se llevó a cabo un análisis de la unión de otros reguladores de *Gai3* mediante RMN, incluyendo los ligandos naturales  $G\beta 1\gamma 2$  y el fragmento GoLoco de RGS12, así como los ligandos no naturales R6A, R6A-1 y KB-1753. Después de analizar los CSPs, se determinó que todos los ligandos se unen al dominio G de *Gai3*, a excepción de RGS12, que también se une al dominio helicoidal. Estos hallazgos son coherentes con las estructuras cristalográficas de homólogos de *Gai3*, con la excepción de  $G\beta 1\gamma 2$ , donde el dominio N-terminal apenas muestra perturbaciones, en contraste con la estructura de *Gai1* (rata)- $G\beta 1\gamma 2$ , en la que se observa un contacto muy estrecho entre esta hélice y  $G\beta$ .

Se realizó un análisis análogo para evaluar el impacto de IGGi-11 en estos complejos. Se agregó un gran exceso del compuesto y se examinaron los CSPs. Se constató que IGGi-11 no tenía la capacidad de desplazar a ninguno de los ligandos, pero sí inducía alteraciones en los complejos, sugiriendo la formación de un complejo ternario, excepto en los casos en los que *Gai3* se encontraba ligada a RGS12 o a KB-1753 (específico para GTP-*Gai3*).

*Gai3* posee tres triptófanos cuyas señales de cadena lateral aparecen en una región aislada del espectro. Esta característica, combinada con su ubicación en la estructura de *Gai3*, los convierte en herramientas valiosas para una rápida evaluación de la unión de ligandos. Estos triptófanos son W131 (en el dominio helicoidal), W211 (en la región switch II) y W258 (en el bucle  $\alpha 3\text{-}\beta 5$ ). En todos los complejos estudiados, se observa una perturbación en las cadenas laterales de W211 y W258, mientras que W131 solo se ve afectado en el complejo de *Gai3* con RGS12.

Se empleó la perturbación de dos de estas señales, W211 y W258, para medir la velocidad de intercambio entre GDP y GTP en *Gai3* mediante RMN. Esto se logró al medir la intensidad de las señales correspondientes a GDP-*Gai3* y GTP-*Gai3* después de agregar un exceso de GTP $\gamma$ S a lo largo del tiempo. Este análisis reveló que la constante de intercambio de *Gai3* es de  $(5.5 \pm 0.7) \cdot 10^{-5} \text{ s}^{-1}$  y se puede emplear para el análisis del impacto de los ligandos en esta cinética de intercambio.



Las proteínas G desempeñan un papel fundamental en la señalización celular al alternar entre dos estados: uno inactivo unido a GDP y otro activo unido a GTP. En su estado activo, estas proteínas regulan una amplia variedad de procesos biológicos. La transición hacia el estado activo implica el intercambio de nucleótidos, mientras que el regreso al estado inactivo se logra mediante la hidrólisis del GTP. A pesar de que las proteínas G tienen la capacidad de llevar a cabo este ciclo por sí solas, suelen contar con una serie de reguladores debido a que las tasas de intercambio y de hidrólisis son lentas. Entre ellos se encuentran los GEF, que aceleran el intercambio de nucleótidos; los GAP, que aceleran la hidrólisis de GTP; y los GDI, que ralentizan la disociación del GDP. Además, según su estructura oligomérica, las proteínas G se dividen en dos categorías: monoméricas y heterotriméricas.

Las proteínas G heterotriméricas se componen de tres subunidades:  $\alpha$ ,  $\beta$  y  $\gamma$ , con la subunidad  $\alpha$  siendo la que se asocia con el nucleótido. Cuando la subunidad  $\alpha$  cambia de GDP a GTP, se disocia del complejo  $\beta\gamma$ , y ambas partes inician sus respectivas cascadas de señalización. Los GEF canónicos para las proteínas G heterotriméricas son los receptores acoplados a proteínas G (GPCRs), que aceleran el intercambio de nucleótidos de guanina en la subunidad  $\alpha$ . Así, las proteínas G heterotriméricas son componentes clave en la transducción de señales desde el exterior de la célula a su interior, a través de la membrana citoplasmática. Sin embargo, algunas de estas proteínas G también pueden ser activadas por proteínas citosólicas.

Una de estas proteínas GEF no receptora es la proteína GIV, que tiene la capacidad de activar  $G\alpha i3$ . Esto desencadena una señalización intracelular que influye en la migración celular, un proceso esencial para la cicatrización de heridas, la respuesta inmunológica y el desarrollo embrionario.

La interacción entre GIV y  $G\alpha i3$  ha sido asociada con cánceres metastásicos. La elevada concentración de GIV en estos tipos de cáncer puede conducir a una regulación desequilibrada de la migración celular y la invasión de tejidos, lo que facilita la diseminación de células cancerosas a otras partes del cuerpo. Por lo tanto, se ha identificado la interrupción de la interacción entre GIV y  $G\alpha i3$  como un posible diana terapéutico para el tratamiento de ciertos tipos de cáncer.

La interacción entre estas moléculas se ha analizado previamente mediante resonancia magnética nuclear (RMN), y se ha demostrado su susceptibilidad a la inhibición farmacológica a través de dos estrategias de detección complementarias. En conjunto,

se lograron identificar ocho compuestos únicos a partir de ambos enfoques. Se realizaron curvas de inhibición en función de la dosis para los ocho compuestos utilizando FP, y se encontró que NF023 presentaba una inhibición dependiente de la dosis con valores de IC<sub>50</sub> cercanos a 5 μM. No obstante, no resulta ser un candidato adecuado para el desarrollo de fármacos, ya que no atraviesa la membrana celular y carece de especificidad. A pesar de esto, estos hallazgos respaldan la posibilidad de dirigirse específicamente a la interacción Gai3-GIV.

Tras este avance, nuestros colaboradores llevaron a cabo un cribado de 200,000 compuestos con el objetivo de identificar un posible inhibidor. En este proceso de cribado, se consideraron diversos factores, como la afinidad de unión a Gai3, la disponibilidad para su adquisición y la viabilidad celular. Entre todos los compuestos analizados, únicamente uno, IGGi-11, demostró tener actividad inhibitoria contra la interacción GIV-Gai3.

En el marco de esta investigación, se llevó a cabo un análisis estructural exhaustivo de la interacción entre IGGi-11 y Gai3, así como de la interacción entre el péptido GIV<sub>1671-1705</sub>, que contiene el motivo GBA de GIV, y Gai3. Dado que Gai3 tiene un peso molecular de 41 kDa, se enriqueció la muestra con deuterio y <sup>15</sup>N de GDP-Gai3. Se realizó una titulación de esta muestra con IGGi-11 y se comparó con el espectro de la interacción entre Gai3 y GIV<sub>1671-1705</sub>.

Los resultados de esta titulación revelaron cambios en los desplazamientos químicos en Gai3 en respuesta a la presencia de IGGi-11, y estos cambios fueron similares a los observados con GIV<sub>1671-1705</sub>. Estos hallazgos sugieren que IGGi-11 actúa como un inhibidor competitivo (no alostérico) en la interacción entre Gai3 y GIV<sub>1671-1705</sub>.

El análisis de la titulación proporcionó una estimación de la constante de disociación, que fue de aproximadamente 6 μM para la interacción Gai3-IGGi-11, una cifra respaldada por resultados similares obtenidos mediante calorimetría.

Dado que GIV se une específicamente a la forma inactiva Gai3, se llevó a cabo una titulación por RMN, similar a la mencionada anteriormente, con el propósito de investigar la interacción entre IGGi-11 y GTP-Gai3. En este caso, las perturbaciones en las señales de RMN causadas por la interacción de IGGi-11 con GTP-Gai3 fueron pequeñas y afectaron las mismas regiones que en el caso de GDP-Gai3. Esto indica que IGGi-11 se une al mismo sitio de unión, pero con una afinidad significativamente



más baja. La afinidad se estimó en  $1014 \pm 235 \mu\text{M}$  a  $30 \text{ }^\circ\text{C}$ , lo que representa una disminución de 188 veces en comparación con la afinidad observada en la interacción con GDP-G $\alpha$ i3. Por lo tanto, IGGi-11 es también específico a la forma inactiva de G $\alpha$ i3.

Se investigó cómo IGGi-11 afecta a las vías de señalización celular reguladas por la interacción G $\alpha$ i3-GIV. Brevemente, tras el estímulo con EGF, GIV se enlaza a EGFR formando el complejo GIV-EGFR junto a G $\alpha$ i3. En células con GIV funcional, esto activa vías de migración y proliferación. Entre ellas, se activa la vía PI3K/Akt, lo que conlleva la activación de proteínas como Akt y sus sustratos, por ejemplo PRAS40. Para analizar el impacto de IGGi-11 en esta cascada de señalización, se emplearon células MDA-MB-231, derivadas del cáncer de mama, que tienen una alta expresión de GIV. Tras el tratamiento con IGGi-11, se observó una reducción en la fosforilación de Akt y PRAS40, lo que sugiere que la cascada de señalización G $\alpha$ i3-GIV se inhibe en presencia del compuesto. Es importante destacar que este efecto no afectó la cantidad total de proteína, lo que indica la especificidad de la acción de IGGi-11 en esta interacción.

La obtención de una estructura cristalina de G $\alpha$ i3 en complejo con IGGi-11 o un péptido GIV sería fundamental para mejorar nuestra comprensión de sus interacciones moleculares. Estas estructuras proporcionarían detalles esenciales sobre el modo de unión y los eventos clave entre G $\alpha$ i3 y el inhibidor o péptido. Esta información permitiría diseñar y optimizar el inhibidor de manera racional, ajustando su afinidad y especificidad. Sin embargo, no ha sido posible cristalizar G $\alpha$ i3 en forma libre o en complejo con ligandos pequeños.

Nuestra hipótesis se centró en que dicha dificultad se debía a la presencia de regiones flexibles que obstaculizaban la cristalización. Por lo tanto, recurrimos a ensayos de relajación mediante RMN de  $^{15}\text{N}$  con el fin de identificar estas áreas flexibles y diseñar una versión más propicia para la cristalización de G $\alpha$ i3 humana. El análisis mediante RMN determinó la existencia de 31 residuos en el extremo N-terminal y 6 residuos en el extremo C-terminal que presentaban alta flexibilidad. En consecuencia, diseñamos la construcción sG $\alpha$ i3, eliminando estas zonas flexibles (G $\alpha$ i3<sub>31-350</sub>) e incorporando una fusión con ubiquitina con una cola de histidinas, lo que facilitó su purificación, aumentó la expresión y mejoró la solubilidad. La ubiquitina se separaba de la proteína mediante un sitio de corte para la proteasa TEV.

Por SEC-MALS se confirmó que sGai3 se encontraba en su forma monomérica con un peso molecular de 37.5 kDa. Mediante ITC a 25 °C, se evaluó la afinidad de IGGi-11 y GIV<sub>1672-1693-DY</sub> por sGai3 y se observó que su afinidad era comparable a la que se encontraba con Gai3 de longitud completa. Tras realizar estos ensayos para asegurarnos de que la construcción era adecuada para los experimentos de cristalización, procedimos con los ensayos de cristalización, tanto con sGai3 en su forma libre como en complejo con GIV<sub>1672-1693-DY</sub> e IGGi-11.

Los ensayos en los que la proteína se encontraba en presencia de un ligando tuvieron éxito en la cristalización, pero no logramos obtener cristales con sGai3 en su forma libre. Tras múltiples etapas de optimización, obtuvimos una densidad electrónica de baja resolución a partir de los cristales de los ensayos en los que sGai3 estaba en presencia de IGGi-11. Al analizar los datos de difracción de estos cristales, nos sorprendió descubrir que proporcionaron la estructura de sGai3 en su forma libre con una resolución de 3.34 Å.

Ante las dificultades en la cristalización, optamos por un enfoque de modelado molecular utilizando HADDOCK. Inicialmente, aplicamos restricciones de distancia basadas en CSPs de la interacción entre IGGi-11 y Gai3. Sin embargo, esta metodología generó resultados ambiguos. Ante la falta de conclusiones claras, se llevó a cabo un docking basado en energías sin restricciones.

Utilizamos SiteMap para identificar dos posibles sitios de unión para los ligandos: uno en la hendidura formada por switch II, la hélice  $\alpha 3$  y la región N-terminal del bucle  $\alpha 3$ - $\beta 5$  de sGai3, y otro en una hendidura en el lado opuesto de la proteína, formada por el bucle entre las hélices C y D, la hélice E del dominio helicoidal y la primera mitad de la región  $\beta 5$ - $\alpha 4$  (a menudo denominada hélice F del dominio helicoidal). Luego, mediante AutoDock, identificamos cinco posibles sitios de unión, pero excluimos cuatro de ellos debido a inconsistencias con datos experimentales de RMN y mutagénesis. Por lo tanto, nos enfocamos en el sitio que involucra switch II, la hélice  $\alpha 3$  y la región N-terminal del bucle  $\alpha 3$ - $\beta 5$  de Gai3, donde encontramos tres poses estables.

Dos de estas poses se agrupan en lo que denominamos "modo 1", que se sitúa más cerca del bucle  $\alpha 3$ - $\beta 5$ , mientras que la tercera se ubica en el "modo 2", donde el núcleo hidrofóbico de fluoreno de IGGi-11 queda más próximo a W211. En ambos modos, el núcleo hidrofóbico de fluoreno de IGGi-11 se encuentra encapsulado en un

bolsillo hidrofóbico de *Gai3* y los grupos carboxilo interactúan con residuos cargados positivamente (R208 en el modo 1 y K210 en el modo 2). Las moléculas acopladas sugieren que modificar la simetría podría estabilizar una de las posibles conformaciones de unión.

Además, se llevó a cabo un análisis de la unión de otros reguladores de *Gai3* mediante RMN, incluyendo los ligandos naturales  $G\beta 1\gamma 2$  y el fragmento GoLoco de RGS12, así como los ligandos no naturales R6A, R6A-1 y KB-1753. Después de analizar los CSPs, se determinó que todos los ligandos se unen al dominio G de *Gai3*, a excepción de RGS12, que también se une al dominio helicoidal. Estos hallazgos son coherentes con las estructuras cristalográficas de homólogos de *Gai3*, con la excepción de  $G\beta 1\gamma 2$ , donde el dominio N-terminal apenas muestra perturbaciones, en contraste con la estructura de *Gai1* (rata)- $G\beta 1\gamma 2$ , en la que se observa un contacto muy estrecho entre esta hélice y  $G\beta$ .

Se realizó un análisis análogo para evaluar el impacto de IGGi-11 en estos complejos. Se agregó un gran exceso del compuesto y se examinaron los CSPs. Se constató que IGGi-11 no tenía la capacidad de desplazar a ninguno de los ligandos, pero sí inducía alteraciones en los complejos, sugiriendo la formación de un complejo ternario, excepto en los casos en los que *Gai3* se encontraba ligada a RGS12 o a KB-1753 (específico para GTP-*Gai3*).

*Gai3* posee tres triptófanos cuyas señales de cadena lateral aparecen en una región aislada del espectro. Esta característica, combinada con su ubicación en la estructura de *Gai3*, los convierte en herramientas valiosas para una rápida evaluación de la unión de ligandos. Estos triptófanos son W131 (en el dominio helicoidal), W211 (en la región switch II) y W258 (en el bucle  $\alpha 3\text{-}\beta 5$ ). En todos los complejos estudiados, se observa una perturbación en las cadenas laterales de W211 y W258, mientras que W131 solo se ve afectado en el complejo de *Gai3* con RGS12.

Se empleó la perturbación de dos de estas señales, W211 y W258, para medir la velocidad de intercambio entre GDP y GTP en *Gai3* mediante RMN. Esto se logró al medir la intensidad de las señales correspondientes a GDP-*Gai3* y GTP-*Gai3* después de agregar un exceso de GTP $\gamma$ S a lo largo del tiempo. Este análisis reveló que la constante de intercambio de *Gai3* es de  $(5.5 \pm 0.7) \cdot 10^{-5} \text{ s}^{-1}$  y se puede emplear para el análisis del impacto de los ligandos en esta tasa de intercambio.





# Bibliography



- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwillinger, T. C., & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, 66(2), 213-221. doi: 10.1107/S0907444909052925
- Anthis, N. J., & Clore, G. M. (2013). Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein science*, 22(6), 851-858. doi: 10.1002/pro.2253
- Aznar, N., Midde, K. K., Dunkel, Y., Lopez-Sanchez, I., Pavlova, Y., Marivin, A., Barbazán, J., Murray, F., Nitsche, U., Janssen, K.-P., Willert, K., Goel, A., Abal, M., Garcia-Marcos, M., & Ghosh, P. (2015). Daple is a novel non-receptor GEF required for trimeric G protein activation in Wnt signaling. *eLife*, 4, e07091. doi: 10.7554/eLife.07091
- Berjanskii, M. V., & Wishart, D. S. (2005). A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. *Journal of the American Chemical Society*, 127(43), 14970–14971. doi: 10.1021/ja054842f
- Bos, J. L., Rehmann, H., & Wittinghofer, A. (2007). GEFs and GAPs: Critical Elements in the Control of Small G Proteins. *Cell*, 129(5), 865-877. doi: 10.1016/j.cell.2007.05.018
- Brüschweiler, R., Liao, X., & Wright, P. E. (1995). Long-range motional restrictions in a multidomain zinc-finger protein from anisotropic tumbling. *Science (New York, N.Y.)*, 268(5212), 886–889. doi: 10.1126/science.7754375
- Case, D. A., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D., Cheatham III, T. E., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Giambasu, G., Gilson, M. K., Gohlke, H., Goetz, A. W., Harris, R., Izadi, S., Izmailov, S. A., Kasavajhala, K., Kovalenko, A., Krasny, R., ... Kollman, P. A. (2020). AMBER 2020, University of California, San Francisco.
- Cavanagh, J., Skelton, N., Palmer, III, A. G., Fairbrother, W. J., & Rance, M. (2010). *Protein NMR Spectroscopy: Principles and Practice*. Academic Press.
- Chen, V. B., Arendall 3rd, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., & Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography*, 66(1), 12-21. doi: 10.1107/S0907444909042073
- Clapman, D. E. (1997). G protein  $\beta\gamma$  subunits. *Annual Review of Pharmacology and Toxicology*, 37(1), 167-203. doi: 10.1146/annurev.pharmtox.37.1.167

- Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G., & Sprang, S. R. (1994). Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. *Science (New York, N.Y.)*, 265(5177), 1405-1412. doi: 10.1126/science.8073283
- Cole, R., & Loria, J. P. (2003). FAST-Modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data. *Journal of biomolecular NMR*, 26(3), 203–213. doi: 10.1023/a:1023808801134
- de Oliveira, P. G., Ramos, M. L. S., Amaro, A. J., Dias, R. A., & Vieira, S. I. (2019). Gi/o-Protein Coupled Receptors in the Aging Brain. *Frontiers in Aging Neuroscience*, 11(89). doi: 10.3389/fnagi.2019.00089
- de Opakua, A. I., Parag-Sharma, K., DiGiacomo, V., Merino, N., Leyme, A., Marivin, A., Villate, M., Nguyen, L. T., de la Cruz-Morcillo, M. Á., Blanco-Canosa, J. B., Ramachandran, S., Baillie, G. S., Cerione, R. A., Blanco, F. J., & Garcia-Marcos, M. (2017). Molecular mechanism of Gai activation by non-GPCR proteins with a Gα-Binding and Activating motif. *Nature communications*, 8(15163). doi: 10.1038/ncomms15163
- DiGiacomo, V., de Opakua, A. I., Papakonstantinou, M. P., Nguyen, L. T., Merino, N., Blanco-Canosa, J. B., Blanco, F. J., & Garcia-Marcos, M. (2017). The Gai-GIV binding interface is a druggable protein-protein interaction. *Scientific reports*, 7(1), 8575. doi: 10.1038/s41598-017-08829-7
- Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, 125(7), 1731–1737. doi: 10.1021/ja026939x
- Dunkel, Y., Ong, A., Notani, D., Mittal, Y., Lam, M., Lam, X., & Ghosh, P. (2012). STAT3 protein up-regulates Gα-interacting vesicle-associated protein (GIV)/Girdin expression, and GIV enhances STAT3 activation in a positive feedback loop during wound healing and tumor invasion/metastasis. *The Journal of biological chemistry*, 287(50), 41667–41683. doi: 10.1074/jbc.M112.390781
- Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. D. (2010). Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography*, 66(4), 486-501. doi: 10.1107/S0907444910007493
- Evans, P. (2006). Scaling and assessment of data quality. *Acta crystallographica. Section D, Biological crystallography*, 62(1), 72-82. doi: 10.1107/S0907444905036693
- Evans, P. R., & Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography*, 69(7), 1204-1214. doi: 10.1107/S0907444913000061



- 
- Ferré, S. (2015). The GPCR heterotetramer: challenging classical pharmacology. *Trends in pharmacological sciences*, 36(3), 145-152. doi: 10.1016/j.tips.2015.01.002
- Flock, T., Hauser, A. S., Lund, N., Gloriam, D. E., Balaji, S., & Babu, M. M. (2017). Selectivity determinants of GPCR-G-protein binding. *Nature*, 545(7654), 317-322. doi: 10.1038/nature22070
- Folta-Stogniew, E., & Williams, K. R. (1999). Determination of Molecular Masses of Proteins in Solution: Implementation of an HPLC Size Exclusion Chromatography and Laser Light Scattering Service in a Core Laboratory. *Journal of biomolecular techniques*, 10(2), 51-63.
- Garcia-Marcos, M. (2023). Stickier G-protein conformations. *Nature Chemical Biology*, 19(6), 665-666. doi: 10.1038/s41589-022-01239-5
- Garcia-Marcos, M., Ghosh, P., & Farquhar, M. G. (2009). GIV is a nonreceptor GEF for Gαi with a unique motif that regulates Akt signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 106(9), 3178–3183. doi: 10.1073/pnas.0900294106
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*, 31(13), 3784–3788. doi: 10.1093/nar/gkg563
- Gebregiworgis, T., Marshall, C. B., Nishikawa, T., Radulovich, N., Sandí, M.-J., Fang, Z., Rottapel, R., Tsao, M.-S., & Ikura, M. (2018). Multiplexed Real-Time NMR GTPase Assay for Simultaneous Monitoring of Multiple Guanine Nucleotide Exchange Factor Activities from Human Cancer Cells and Organoids. *Journal of the American Chemical Society*, 140(13), 4473-4476. doi: 10.1021/jacs.7b13703
- Geyer, M., & Wittinghofer, A. (1997). GEFs, GAPs, GDIs and effectors: taking a closer (3D) look at the regulation of Ras-related GTP-binding proteins. *Current Opinion in Structural Biology*, 7(6), 786-792. doi: 10.1016/S0959-440X(97)80147-9
- Ghosh, P., Beas, A. O., Bornheimer, S. J., Garcia-Marcos, M., Forry, E. P., Johansson, C., Ear, J., Jung, B. H., Cabrera, B., Carethers, J. M., & Farquhar, M. G. (2010). A Gαi-GIV molecular complex binds epidermal growth factor receptor and determines whether cells migrate or proliferate. *Molecular biology of the cell*, 21(13), 2338-2354. doi: 10.1091/mbc.E10-01-0028
- Ghosh, P., Garcia-Marcos, M., Bornheimer, S. J., & Farquhar, M. G. (2008). Activation of Gαi3 triggers cell migration via regulation of GIV. *The Journal of cell biology*, 182(2), 381-393. doi: 10.1083/jcb.200712066
- Ghosh, P., Beas, A. O., Bornheimer, S. J., Garcia-Marcos, M., Forry, E. P., Johansson, C., Ear, J., Jung, B. H., Cabrera, B., Carethers, J. M., & Farquhar, M. G. (2010). A Gαi-GIV Molecular Complex Binds Epidermal Growth Factor Receptor and
-

- Determines Whether Cells Migrate or Proliferate. *Molecular biology of the cell*, 21(13), 2338–2354. doi: 10.1091/mbc.e10-01-0028
- Ghosh, P., Rangamani, P., & Kufareva, I. (2017). The GAPs, GEFs, GDIs and...now, GEMs: New kids on the heterotrimeric G protein signaling block. *Cell cycle (Georgetown, Tex.)*, 16(7), 607-612. doi: 10.1080/15384101.2017.1282584
- Goto, S. (2017). Striatal G $\alpha$ olf/cAMP Signal-Dependent Mechanism to Generate Levodopa-Induced Dyskinesia in Parkinson's Disease. *Frontiers in cellular neuroscience*, 11(364). doi: 10.3389/fncel.2017.00364
- Guo, P., Tai, Y., Wang, M., Sun, H., Zhang, L., Wei, W., Xiang, Y. K., & Wang, Q. (2022). G $\alpha$ 12 and G $\alpha$ 13: Versatility in Physiology and Pathology. *Frontiers in cell and developmental biology*, 10(809425). doi: 10.3389/fcell.2022.809425
- Gupta, V., Bhandari, D., Leyme, A., Aznar, N., Midde, K. K., Lo, I.-C., Ear, J., Niesman, I., López-Sánchez, I., Blanco-Canosa, J. B., von Zastrow, M., Garcia-Marcos, M., Farquhar, M. G., & Ghosh, P. (2016). GIV/Girdin activates Gai and inhibits Gas via the same motif. *Proceedings of the National Academy of Sciences of the United States of America*, 113(39), E5721–E5730. doi: 10.1073/pnas.1609502113
- Halgren, T. A. (2009). Identifying and characterizing binding sites and assessing druggability. *Journal of chemical information and modeling*, 49(2), 377-389. doi: 10.1021/ci800324m
- He, Q., Bouley, R., Liu, Z., Wein, M. N., Zhu, Y., Spatz, J. M., Wang, C.-Y., Divieti Pajevic, P., Plagge, A., Babitt, J. L., & Bastepe, M. (2017). Large G protein  $\alpha$ -subunit XL $\alpha$ s limits clathrin-mediated endocytosis and regulates tissue iron levels in vivo. *Proceedings of the National Academy of Sciences*, 114(45), 9559-9568. doi: 10.1073/pnas.1712670114
- Hinton, D. R., Blanks, J. C., Fong, H. K. W., Casey, P. J., Hildebrandt, E., & Simons, M. I. (1990). Novel localization of a G protein, Gz-alpha, in neurons of brain and retina. *Journal of Neuroscience*, 10(8), 2763-2770. doi:10.1523/JNEUROSCI.10-08-02763.1990
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of molecular graphics*, 14(1), 33-38. doi: 10.1016/0263-7855(96)00018-5
- Ivanina, T., Varon, D., Peleg, S., Rishal, I., Porozov, Y., Dessauer, C. W., Keren-Raifman, T., & Dascal, N. (2004). Gai1 and Gai3 Differentially Interact with, and Regulate, the G Protein-activated K<sup>+</sup> Channel. *The Journal of biological chemistry*, 279(17), 17260–17268. doi: 10.1074/jbc.M313425200
- Ja, W. W., Adhikari, A., Austin, R. J., Sprang, S. R., & Roberts, R. W. (2005). A peptide core motif for binding to heterotrimeric G protein alpha subunits. *The Journal of biological chemistry*, 280(37), 32057–32060. doi: 10.1074/jbc.C500319200

- 
- Ja, W. W., & Roberts, R. W. (2004). In Vitro Selection of State-Specific Peptide Modulators of G Protein Signaling Using mRNA Display. *Biochemistry*, 43(28), 9265–9275. doi: 10.1021/bi0498398
- Johnston, C. A., Lobanova, E. S., Shavkunov, A. S., Low, J., Ramer, J. K., Blaesius, R., Fredericks, Z., Willard, F. S., Kuhlman, B., Arshavsky, V. Y., & Siderovski, D. P. (2006). Minimal determinants for binding activated G alpha from the structure of a G alpha(i1)-peptide dimer. *Biochemistry*, 45(38), 11390-11400. doi: 10.1021/bi0613832
- Johnston, C. A., Willard, F. S., Jezyk, M. R., Fredericks, Z., Bodor, E. T., Jones, M. B., Blaesius, R., Watts, V. J., Harden, T. K., Sondek, J., Ramer, J. K., & Siderovski, D. P. (2005). Structure of Galpha(i1) bound to a GDP-selective peptide provides insight into guanine nucleotide exchange. *Structure (London, England : 1993)*, 13(7), 1069–1080. doi: 10.1016/j.str.2005.04.007
- Johnston, C. A., Willard, F. S., Ramer, J. K., Blaesius, R., Roques, C. N., & Siderovski, D. P. (2008). State-selective Binding Peptides for Heterotrimeric G-protein Subunits: Novel Tools for Investigating G-protein Signaling Dynamics. *Combinatorial chemistry & high throughput screening*, 11(5), 370-381. doi: 10.2174/138620708784534798
- Jose, J., & Choudhury, S. R. (2020). Heterotrimeric G-proteins mediated hormonal responses in plants. *Cellular signalling*, 76(109799). doi: 10.1016/j.cellsig.2020.109799
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. doi: 10.1038/s41586-021-03819-2
- Kabsch, W. (2010). XDS. *Acta Crystallographica Section D: Biological Crystallography*, 66(2), 125-132. doi: 10.1107/S0907444909047337
- Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12), 2577–2637. doi: 10.1002/bip.360221211
- Kalogriopoulos, N. A., Rees, S. D., Ngo, T., Kopcho, N. J., Ilatovskiy, A. V., Sun, N., Komives, E. A., Chang, G., Ghosh, P., & Kufareva, I. (2019). Structural basis for GPCR-independent activation of heterotrimeric Gi proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 116(33), 16394–16403. doi: 10.1073/pnas.1906658116
- Kamato, D., Mitra, P., Davis, F., Osman, N., Chaplin, R., Cabot, P. J., Afroz, R., Thomas, W., Zheng, W., Kaur, H., Brimble, M., & Little, P. J. (2017). Gaq proteins: molecular
-

- pharmacology and therapeutic potential. *Cellular and Molecular Life Sciences*, 74(8), 1379–1390. doi: 10.1007/s00018-016-2405-9
- Kamato, D., Thach, L., Bernard, R., Chan, V., Zheng, W., Kaur, H., Brimble, M., Osman, N., & Little, P. J. (2015). Structure, function, pharmacology, and therapeutic potential of the G protein,  $G\alpha_q11$ . *Frontiers in cardiovascular medicine*, 2(14). doi: 10.3389/fcvm.2015.00014
- Kimple, A. J., Soundararajan, M., Hutsell, S. Q., Roos, A. K., Urban, D. J., Setola, V., Temple, B. R. S., Roth, B. L., Knapp, S., Willard, F. S., & Siderovski, D. P. (2009). Structural determinants of G-protein alpha subunit selectivity by regulator of G-protein signaling 2 (RGS2). *The Journal of biological chemistry*, 284(29), 19402–19411. doi: 10.1074/jbc.M109.024711
- Kimple, R. J., De Vries, L., Tronchère, H., Behe, C. I., Morris, R. A., Farquhar, M. G., & Siderovski, D. P. (2001). RGS12 and RGS14 GoLoco motifs are  $G\alpha(i)$  interaction sites with guanine nucleotide dissociation inhibitor. *The Journal of biological chemistry*, 276(31), 29275–29281. doi: 10.1074/jbc.M103208200
- Kimple, R. J., Kimple, M. E., Betts, L., Sondek, J., & Siderovski, D. P. (2002). Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. *Nature*, 416(6883), 878–881. doi: 10.1038/416878a
- Kleuss, C., Raw, A. S., Lee, E., Sprang, S. R., & Gilman, A. G. (1994). Mechanism of GTP hydrolysis by G-protein alpha subunits. *Proceedings of the National Academy of Sciences of the United States of America*, 91(21), 9828–9831. doi: 10.1073/pnas.91.21.9828
- Kobilka, B. K. (2007). G protein coupled receptor structure and activation. *Biochimica et Biophysica Acta*, 1768(4), 794–807. doi: 10.1016/j.bbamem.2006.10.021
- Lee, W., Tonelli, M., & Markley, J. L. (2015). NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics (Oxford, England)*, 31(8), 1325–1327. doi: 10.1093/bioinformatics/btu830
- Lerea, C. L., Somers, D. E., Hurley, J. B., Klock, I. B., & Bunt-Milam, A. H. (1986). Identification of specific transducin alpha subunits in retinal rod and cone photoreceptors. *Science*, 234(4772), 77–80. doi: 10.1126/science.3529395
- Li, J., Ge, Y., Huang, J.-X., Strømgaard, K., Zhang, X., & Xiong, X.-F. (2020). Heterotrimeric G Proteins as Therapeutic Targets in Drug Discovery. *Journal of medicinal chemistry*, 63(10), 5013–5030. doi: 10.1021/acs.jmedchem.9b01452
- Lin, S., Han, S., Cai, X., Tan, Q., Zhou, K., Wang, D., Wang, X., Du, J., Yi, C., Chu, X., Dai, A., Zhou, Y., Chen, Y., Zhou, Y., Liu, H., Liu, J., Yang, D., Wang, M.-W., Zhao, Q., & Wu, B. (2021). Structures of Gi-bound metabotropic glutamate receptors mGlu2 and mGlu4. *Nature*, 594(7864), 583–588. doi: 10.1038/s41586-021-03495-2

- Liu, W., & Northup, J. K. (1998). The helical domain of a G protein  $\alpha$  subunit is a regulator of its effector. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 12878-12883. doi: 10.1073/pnas.95.22.12878
- Lipari, G., & Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *Journal of the American Chemical Society*, 104(17), 4559-4570. doi: 10.1021/ja00381a010
- Mase, Y., Yokogawa, M., Osawa, M., & Shimada, I. (2012). Backbone resonance assignments for G protein  $\alpha(i3)$  subunit in the GTP-bound state. *Biomolecular NMR assignments*, 6(2), 217-220. doi: 10.1007/s12104-012-9361-6
- Mase, Y., Yokogawa, M., Osawa, M., & Shimada, I. (2014). Backbone resonance assignments for G protein  $\alpha i3$  subunit in the GDP-bound state. *Biomolecular NMR Assignments*, 8(2), 237-241. doi: 10.1007/s12104-013-9491-5
- Mazhab-Jafari, M. T., Marshall, C. B., Smith, M., Gasmi-Seabrook, G. M. C., Stambolic, V., Rottapel, R., Neel, B. G., & Ikura, M. (2010). Real-time NMR study of three small GTPases reveals that fluorescent 2'(3')-O-(N-methylanthraniloyl)-tagged nucleotides alter hydrolysis and exchange kinetics. *The Journal of biological chemistry*, 285(8), 5132–5136. doi: 10.1074/jbc.C109.064766
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography*, 40(4), 658-674. doi: 10.1107/S0021889807021206
- McEwen, D. P., Gee, K. R., Kang, H. C., & Neubig, R. R. (2001). Fluorescent BODIPY-GTP analogs: real-time measurement of nucleotide binding to G proteins. *Analytical biochemistry*, 291(1), 109-117. doi: 10.1006/abio.2001.5011
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P., & Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research*, 41(Web Server issue), W597–W600. doi: 10.1093/nar/gkt376
- Mlynárik, V. (2017). Introduction to nuclear magnetic resonance. *Analytical biochemistry*, 529, 4-9. doi: 10.1016/j.ab.2016.05.006
- Morris, G. M., Huey, R., Lindstorm, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16), 2785–2791. doi: 10.1002/jcc.21256
- Nguyen, N. T., Nguyen, T. H., Pham, T. N. H., Huy, N. T., Bay, M. V., Pham, M. Q., Nam, P. C., Vu, V. V., & Ngo, S. T. (2020). Autodock Vina Adopts More Accurate Binding Poses but Autodock4 Forms Better Binding Affinity. *Journal of chemical information and modeling*, 60(1), 204-211. doi: 10.1021/acs.jcim.9b00778

- Northup, J. K., Sternweis, P. C., Smigel, M. D., Schleifer, L. S., Ross, E. M., & Gilman, A. G. (1980). Purification of the regulatory component of adenylate cyclase. *Proceedings of the National Academy of Sciences*, 77(11), 6516-6520. doi: 10.1073/pnas.77.11.6516
- Orun, O. (2006). A structural approach to g-protein signaling mechanisms:  $\alpha$ -subunits. *Marmara Medical Journal*, 19(1), 41-45.
- Palmer III, A. G. (Columbia University). PDB Inertia.
- Palmer III, A. G. (Columbia University). quadric\_diffusion.
- Palmer III, A. G., Rance, M., & Wright, P. E. (1991). Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance carbon-13 heteronuclear NMR spectroscopy. *Journal of the American Chemical Society*, 113(12), 4371-4380. doi: 10.1021/ja00012a001
- Pawley, N. H., Wang, C., Koide, S., & Nicholson, L. K. (2001). An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of  $^{15}\text{N}$  relaxation parameters. *Journal of biomolecular NMR*, 20(2), 149-165. doi: 10.1023/a:1011249816560
- Pons, T., Gómez, R., China, G., & Valencia, A. (2003). Beta-propellers: associated functions and their role in human diseases. *Current Medicinal Chemistry*, 10(6), 505-524. doi: 10.2174/09298670333368204
- Rogov, V. V., Rozenknop, A., Rogova, N. Y., Löhr, F., Tikole, S., Jaravine, V., Güntert, P., Dikic, I., & Dötsch, V. (2012). A Universal Expression Tag for Structural and Functional Studies of Proteins. *Chembiochem*, 13(7), 959-963. doi: 10.1002/cbic.201200045
- Sarvazyan, N. A., Remmers, A. E., & Neubig, R. R. (1998). Determinants of  $\text{G}_i$   $\alpha$  and  $\beta$   $\gamma$  binding. Measuring high affinity interactions in a lipid environment using flow cytometry. *The Journal of biological chemistry*, 273(14), 7934-7940. doi: 10.1074/jbc.273.14.7934
- Schrödinger, L., & DeLano, W. L. (2020). PyMOL. Retrieved from <http://www.pymol.org/pymol>.
- Schrödinger, LLC. (2020). Schrödinger Release 2020-2: Maestro. Schrödinger.
- Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., & Uchimaya, M. (2007). Epik: a software program for  $\text{pK}_a$  prediction and protonation state generation for drug-like molecules. *Journal of computer-aided molecular design*, 21(12), 681-691. doi: 10.1007/s10822-007-9133-z
- Siderovski, D. P., Diversé-Pierluissi, M. A., & De Vries, L. (1999). The GoLoco motif: a  $\text{G}\alpha_{\text{hi/o}}$  binding motif and potential guanine-nucleotide exchange factor. *Trends in biochemical sciences*, 24(9), 340-341. doi: 10.1016/s0968-0004(99)01441-3

- 
- Siderovski, D. P., & Willard, F. S. (2005). The GAPs, GEFs, and GDIs of heterotrimeric G-protein alpha subunits. *International Journal of Biological Sciences*, 1(2), 51-66. doi: 10.7150/ijbs.1.51
- Smith, S. J., & Rittinger, K. (2002). Preparation of GTPases for structural and biophysical analysis. *Methods in molecular biology* (Clifton, N.J.), 189, 13–24. doi: 10.1385/1-59259-281-3:013
- Soundararajan, M., Willard, F. S., Kimple, A. J., Turnbull, A. P., Ball, L. J., Schoch, G. A., Gileadi, C., Fedorov, O. Y., Dowler, E. F., Higman, V. A., Hutsell, S. Q., Sundström, M., Doyle, D. A., & Siderovski, D. P. (2008). Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits. *Proceedings of the National Academy of Sciences of the United States of America*, 105(17), 6457-6462. doi: 10.1073/pnas.0801508105
- Spielman, A. I. (1998). Gustducin and its role in taste. *Journal of Dental Research*, 77(4), 539-544. doi: 10.1177/00220345980770040601
- Sprang, S. R. (2016). Invited review: Activation of G Proteins by GTP and the Mechanism of Ga-Catalyzed GTP Hydrolysis. *Biopolymers*, 105(8), 449–462. doi: 10.1002/bip.22836
- Srivastava, D., & Artemyev, N. O. (2020). Ric-8A, a GEF, and a Chaperone for G Protein  $\alpha$ -Subunits: Evidence for the Two-Faced Interface. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 42(3), e1900208. doi: 10.1002/bies.201900208
- Srivastava, V. K., & Yadav, R. (2019). Chapter 9 - Isothermal titration calorimetry. In *Data Processing Handbook for Complex Biological Data Sources* (Gauri Misra ed., pp. 125-137). Elsevier Science. doi: 10.1016/B978-0-12-816548-5.00009-5
- Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein expression and purification*, 41(1), 207-234. doi: 10.1016/j.pep.2005.01.016
- Sternweis, P. C., Northup, J. K., Smigel, M. D., & Gilman, A. G. (1981). The regulatory component of adenylate cyclase. Purification and properties. *The Journal of biological chemistry*, 256(22), 11517–11526.
- Suzuki, N., Hajicek, N., & Kozasa, T. (2009). Regulation and physiological functions of G12/13-mediated signaling pathways. *Neuro-Signals*, 17(1), 55-70. doi: 10.1159/000186690
- Syrovatkina, V., Alegre, K. O., Rey, R., & Huang, X.-Y. (2016). Regulation, Signaling, and Physiological Functions of G-Proteins. *Journal of Molecular Biology*, 428(19), 3850-3868. doi: 10.1016/j.jmb.2016.08.002
- Tennakoon, M., Senarath, K., Kankanamge, D., Ratnayake, K., Wijayarathna, D., Olupothage, K., Ubeyasinghe, S., Martins-Cannavino, K., Hébert, T. E., & Karunarathne, A.
-

- (2021). Subtype-dependent regulation of G $\beta\gamma$  signalling. *Cellular signalling*, 82(109947). doi: 10.1016/j.cellsig.2021.109947
- Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Smart, O., Vonnrhein, C., & Bricogne, G. (2018). STARANISO. Cambridge, United Kingdom: Global Phasing Ltd. <http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>
- Tjandra, N., Feller, S. E., Pastor, R. W., & Bax, A. (1995). Rotational diffusion anisotropy of human ubiquitin from <sup>15</sup>N NMR relaxation. *Journal of the American Chemical Society*, 117(50), 12562-12566. doi: 10.1021/ja00155a020
- Trabalzini, L., & Retta, S. F. (2014). *Ras Signaling: Methods and Protocols*. Humana Press.
- Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2), 455-461. doi: 10.1002/jcc.21334
- Vellano, C. P., Lee, S. E., Dudek, S. M., & Hepler, J. R. (2011). RGS14 at the interface of hippocampal signaling and synaptic plasticity. *Trends in pharmacological sciences*, 32(11), 666-674. doi: 10.1016/j.tips.2011.07.005
- Vieira, T. F., & Sousa, S. F. (2019). Comparing AutoDock and Vina in Ligand/Decoy Discrimination for Virtual Screening. *Applied Sciences*, 9(21), 4538. doi: 10.3390/app9214538
- Vink, R., & Nechifor, M. (Eds.). (2011). *Intracellular magnesium homeostasis*. In: *Magnesium in the Central Nervous System*. University of Adelaide Press. <https://www.ncbi.nlm.nih.gov/books/NBK507258/>
- Vonnrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., & Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. *Acta crystallographica. Section D, Biological crystallography*, 67(4), 293-302. doi: 10.1107/S0907444911007773
- Wedegaertner, P. B., Wilson, P. T., & Bourne, H. R. (1995). Lipid Modifications of Trimeric G Proteins (\*). *Journal of Biological Chemistry*, 270(2), 503-506. doi: 10.1074/jbc.270.2.503
- Wennerberg, K., Rossman, K. L., & Der, C. J. (2005). The Ras superfamily at a glance. *Journal of cell science*, 118(5), 843-6. doi:10.1242/jcs.01660
- Wettschureck, N., & Offermanns, S. (2005). Mammalian G proteins and their cell type specific functions. *Physiological reviews*, 85(4), 1159-1204. doi:10.1152/physrev.00003.2005
- Willard, F. S., & Siderovski, D. P. (2006). The R6A-1 peptide binds to switch II of Galpha11 but is not a GDP-dissociation inhibitor. *Biochemical and biophysical research communications*, 339(4), 1107-1112. doi: 10.1016/j.bbrc.2005.11.132



- 
- Williamson, M. P. (2013). Using chemical shift perturbation to characterise ligand binding. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 73, 1-16. doi: 10.1016/j.pnmrs.2013.02.001
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography*, 67(4), 235-242. doi: 10.1107/S0907444910045749
- Wittinghofer, A., & Vetter, I. R. (2011). Structure-Function Relationships of the G Domain, a Canonical Switch Motif. *Annual Review of Biochemistry*, 80(1), 943-971. doi: 10.1146/annurev-biochem-062708-134043
- Yagi, H., Onoyama, I., Asanoma, K., Hori, E., Yasunaga, M., Kodama, K., Kijima, M., Ohgami, T., Kaneki, E., Okugawa, K., Yahata, H., & Kato, K. (2019). Gal3-mediated LATS1 down-regulation contributes to epithelial-mesenchymal transition in ovarian cancer. *ASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 33(12), 13683–13694. doi: 10.1096/fj.201901278R
- Yang, D., Zhou, Q., Labroska, V., Qin, S., Darbalaei, S., Wu, Y., Yuliantie, E., Xie, L., Tao, H., Cheng, J., Liu, Q., Zhao, S., Shui, W., Jiang, Y., & Wang, M.-W. (2021). G protein-coupled receptors: structure- and function-based drug discovery. *Signal Transduction and Targeted Therapy*, 6(7). doi: 10.1038/s41392-020-00435-w
- Zala, D., Schlattner, U., Desvignes, T., Bobe, J., Roux, A., Chavrier, P., & Boissan, M. (2017). The advantage of channeling nucleotides for very processive functions. *F1000Research*, 6(724). doi: 10.12688/f1000research.11561.2
- Zhao, J., DiGiacomo, V., Ferreras-Gutierrez, M., Dastjerdi, S., Ibáñez de Opakua, A., Park, J.-C., Luebbers, A., Chen, Q., Beeler, A., Blanco, F. J., & Garcia-Marcos, M. (2023). Small-molecule targeting of GPCR-independent non-canonical G protein signaling inhibits cancer progression. *Proceedings of the National Academy of Sciences*. doi: 10.1101/2023.02.18.529092
-





# Publications





# Small-molecule targeting of GPCR-independent noncanonical G-protein signaling in cancer

Jingyi Zhao<sup>a,1</sup> , Vincent DiGiacomo<sup>a,1,2,3</sup> , Mariola Ferreras-Gutierrez<sup>b</sup> , Shiva Dastjerdi<sup>c</sup>, Alain Ibáñez de Opakua<sup>d</sup> , Jong-Chan Park<sup>a</sup>, Alex Luebbbers<sup>a</sup>, Qingyan Chen<sup>a</sup>, Aaron Beeler<sup>c</sup>, Francisco J. Blanco<sup>b</sup> , and Mikel Garcia-Marcos<sup>a,e,3</sup>

Edited by Robert Lefkowitz, HHMI, Durham, NC; received July 31, 2022; accepted March 6, 2023

Activation of heterotrimeric G-proteins ( $G\alpha\beta\gamma$ ) by G-protein-coupled receptors (GPCRs) is a quintessential mechanism of cell signaling widely targeted by clinically approved drugs. However, it has become evident that heterotrimeric G-proteins can also be activated via GPCR-independent mechanisms that remain untapped as pharmacological targets. GIV/Girdin has emerged as a prototypical non-GPCR activator of G proteins that promotes cancer metastasis. Here, we introduce IGGi-11, a first-in-class small-molecule inhibitor of noncanonical activation of heterotrimeric G-protein signaling. IGGi-11 binding to G-protein  $\alpha$ -subunits ( $G\alpha_i$ ) specifically disrupted their engagement with GIV/Girdin, thereby blocking noncanonical G-protein signaling in tumor cells and inhibiting proinvasive traits of metastatic cancer cells. In contrast, IGGi-11 did not interfere with canonical G-protein signaling mechanisms triggered by GPCRs. By revealing that small molecules can selectively disable noncanonical mechanisms of G-protein activation dysregulated in disease, these findings warrant the exploration of therapeutic modalities in G-protein signaling that go beyond targeting GPCRs.

GPCR | G protein | drug discovery | cancer

G-protein-coupled receptors (GPCRs) mediate a large fraction of all transmembrane signaling in the human body, including responses triggered by every major neurotransmitter and by two-thirds of hormones (1). They are also the largest family of druggable proteins in the human genome, representing the target for over one-third of clinically approved drugs (2). To relay signals, GPCRs activate heterotrimeric G-proteins ( $G\alpha\beta\gamma$ ) in the cytoplasm by promoting the exchange of GDP for GTP on  $G\alpha$  subunits, which results in a concomitant dissociation of  $G\beta\gamma$  dimers (3). In turn,  $G\alpha$ -GTP and “free”  $G\beta\gamma$  act on downstream effectors to propagate signaling. Signaling is turned off by the intrinsic GTPase activity of  $G\alpha$ , which leads to the reassociation of  $G\alpha$  with  $G\beta\gamma$ . There is also a growing number of cytoplasmic proteins that modulate nucleotide handling by G-proteins, thereby exerting profound effects on the duration and amplitude of signaling (4–11).

In stark contrast to GPCRs, there are no clinically approved drugs for heterotrimeric G-proteins, despite their well-documented potential as pharmacological targets (12). Small-molecule inhibitors of  $G\beta\gamma$  have been validated in some preclinical models (12, 13), but no drug-like small molecule that targets  $G\alpha$  subunits has been validated. There are, however, some natural cyclic depsipeptides that block  $\alpha$ -subunits of the  $G_{q/11}$  family with high specificity and potency (14). Unfortunately, because they inhibit G-protein activation en toto, these compounds could cause undesired side effects due to indiscriminate blockade of ubiquitous, physiologically relevant functions of their target G-proteins.

Perhaps, a more nuanced targeting approach that exploits disease-specific mechanisms of G-protein regulation could pave the way for new pharmacology. This idea is thwarted by the realization that the mechanisms of G-protein regulation beyond ubiquitous GPCR-mediated activation remain poorly understood in the absence of adequate tools to interrogate them. GIV (also known as Girdin) is a cytoplasmic protein that binds to  $G\alpha_i$  subunits to promote G-protein signaling in a GPCR-independent manner (8, 15–17) and its expression in human primary solid tumors correlates with progression toward more invasive, metastatic stages in various types of cancer (18–20). Tumor cells depleted of GIV also fail to migrate in vitro or metastasize in mice (21). Here, we report the identification of a small molecule that binds to  $G\alpha_i$  to selectively prevent GIV binding without disturbing other mechanisms by which the G-protein is regulated, including canonical GPCR-mediated signaling. We leverage this compound to establish that GIV-mediated activation of G-protein signaling favors proinvasive traits of cancer cells by operating downstream of receptor tyrosine kinases (RTKs) instead of downstream of GPCRs.

## Significance

Human diseases frequently arise from defects in the mechanisms by which external cues are sensed and relayed to the interior of the cell. The proteins most widely targeted by existing therapeutic agents belong to a large family of cell surface receptors named G-protein-coupled receptors (GPCRs), which relay external cues by activating G-proteins in the interior of cells. Here, we report the surprising discovery of a synthetic small molecule that selectively targets G-proteins without compromising their ability to relay signals from GPCRs. Instead, this small molecule disrupts an atypical, GPCR-independent mechanism of G-protein signaling involved in cancer. This work reveals an alternative paradigm in targeting components of a signaling machinery with broad relevance in cellular communication in health and disease.

Author contributions: J.Z., V.D., A.B., F.J.B., and M.G.-M. designed research; J.Z., V.D., M.F.-G., S.D., A.I.d.O., J.-C.P., A.L., and Q.C. performed research; J.Z., V.D., M.F.-G., S.D., A.I.d.O., F.J.B., and M.G.-M. analyzed data; and J.Z., V.D., and M.G.-M. wrote the paper.

Competing interest statement: Boston University has filed a provisional patent application related to the content of this manuscript in which M.G.-M. is listed as an inventor.

This article is a PNAS Direct Submission.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

<sup>1</sup>J.Z. and V.D. contributed equally to this work.

<sup>2</sup>Present address: Jorna Therapeutics, Division of Therapeutic Innovation, Cambridge, MA 02139.

<sup>3</sup>To whom correspondence may be addressed. Email: Vincent.DiGiacomo@gmail.com or mgm1@bu.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2213140120/-DCSupplemental>.

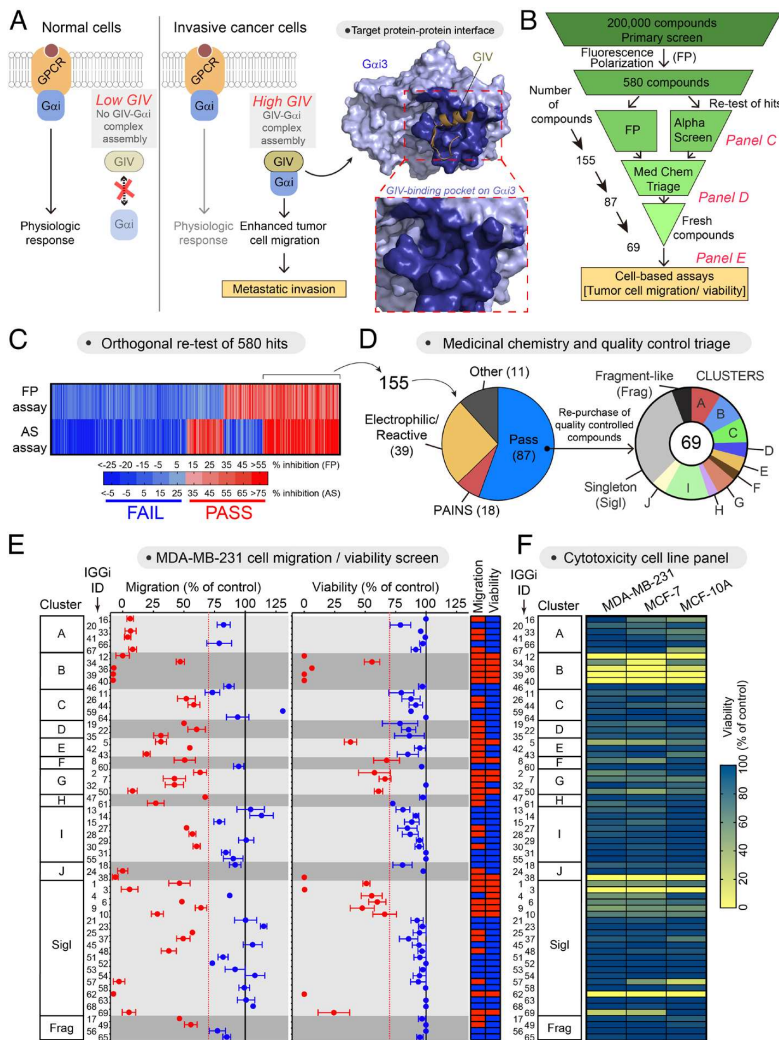
Published April 25, 2023.

## Results

### High-Throughput Screen for Inhibitors of the GIV-G $\alpha$ i Interaction.

Previous work indicates that expression of GIV at high levels in cancer cells might facilitate its association with G $\alpha$ i, which in turn favors tumor cell migration and other prometastatic traits (8, 15–17, 22–25) (Fig. 1A). Moreover, characterization of the molecular basis for the GIV-G $\alpha$ i interaction (Fig. 1A) revealed that this protein-protein interaction might be suitable for specific pharmacological disruption (26–28). These previous findings motivated us to pursue a small-molecule screen for inhibitors of the GIV-G $\alpha$ i interaction. Using a fluorescence polarization (FP) assay that directly monitors GIV binding to G $\alpha$ i3 (27), we obtained 580 hits from screening a collection of 200,000 compounds (Fig. 1B and C). Of these, 155 tested positive for inhibition in both the primary FP assay and an orthogonal secondary assay (AlphaScreen<sup>®</sup>, AS) (27)

(Fig. 1C and D). After triage, 68 compounds were discarded based on unfavorable chemical properties, and only 69 of the remaining 87 compounds could be repurchased as fresh powder stocks (Fig. 1D and *SI Appendix, Table S1*). We named this set of compounds “IGGi,” for “Inhibitors of the GIV-G $\alpha$ i interaction.” We next evaluated the performance of these 69 IGGi compounds in cell-based assays. In cancer cell lines that express high levels of GIV (e.g., the triple-negative metastatic breast cancer cell line, MDA-MB-231), loss of GIV or disruption of its ability to bind G $\alpha$ i through mutagenesis impairs cell migration, but does not affect cell viability under standard in vitro culture conditions on plastic dishes (17, 21). We found that approximately one-third of the IGGi compounds impaired MDA-MB-231 cell migration without affecting viability (Fig. 1E), lending confidence on the ability of our biochemical screen to identify compounds with the desired biological activity. To further prioritize the 69 IGGi compounds, we excluded not



**Fig. 1.** Small-molecule screening to identify inhibitors of the GIV-G $\alpha$ i interaction. (A) Diagram depicting the rationale for targeting the GIV-G $\alpha$ i interaction with small molecules. (B) Scheme of the full screening campaign. (C) Confirmation of hit compounds that inhibit the GIV-G $\alpha$ i interaction in two orthogonal biochemical assays, FP and AS. (D) Triage of compounds based on unfavorable chemical properties and availability of quality controlled molecules. (E) Test of 69 IGGi compounds (100  $\mu$ M) on MDA-MB-231 cell migration and viability. Red, <30% reduction; blue, >30% reduction. Mean  $\pm$  SEM ( $N = 4$ ). (F) Comparison of the effect of IGGi compounds (100  $\mu$ M) on the viability of three breast cell lines, MDA-MB-231, MCF-7, and MCF-10A (mean of  $N = 3$ ).

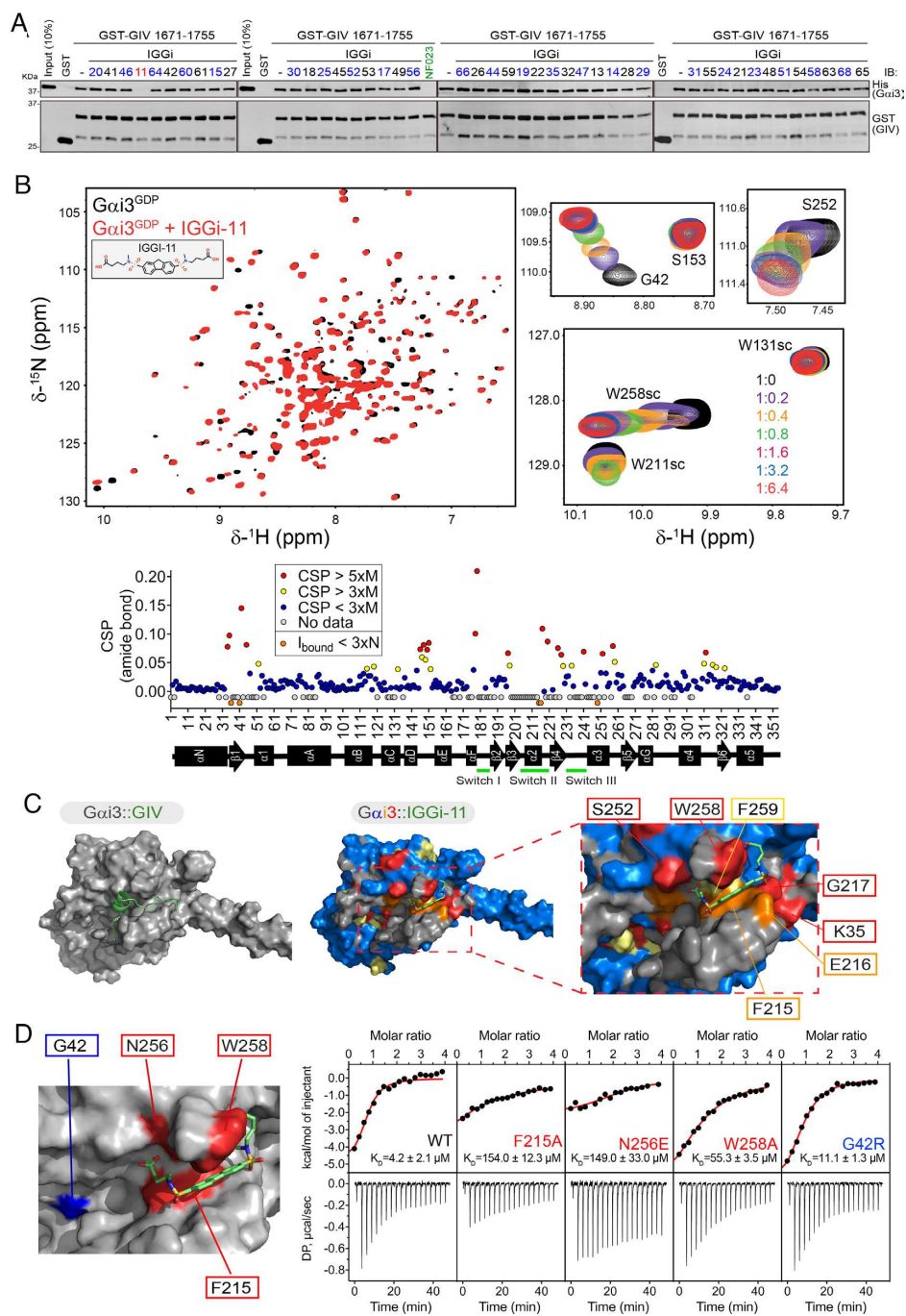
only those with the undesired property of reducing MDA-MB-231 viability, but also those that reduced the viability of MCF-7 cells (a nonmetastatic breast cancer cell line that expresses low levels of GIV) or of MCF-10A (a nontransformed epithelial breast cell line) to eliminate molecules with nonspecific cytotoxicity (Fig. 1*F*). The remaining 44 compounds were tested in a tertiary GIV–G $\alpha$ i binding assay based on GST-fusion pull-downs (PD) (Fig. 2*A*). As a positive control in this assay, we used NF023, a nonselective inhibitor of G $\alpha$ i activity that also disrupts the GIV–G $\alpha$ i interaction in cell-free systems (27, 29). Only one IGGi compound, IGGi-11, was found to inhibit G $\alpha$ i3 binding to GIV in this assay. Despite the weak activity of this compound in MDA-MB-231 cell migration assays (Fig. 1*E*), we pursued its characterization further and experiments presented below indicated high specificity and suitability for cell-based systems upon analog development via relatively minor modification.

**IGGi-11 Binds to the GIV-Interacting Region of G $\alpha$ i.** We reasoned that inhibitors of the GIV–G $\alpha$ i interaction should bind to the G-protein because our primary screening assay used a small peptide fragment of GIV unlikely to harbor enough structural features to accommodate a small molecule. Using NMR spectroscopy, we found that IGGi-11 caused dose-dependent CSPs in the amide bond signals of discrete amino acids of isotopically labeled ( $^2\text{H}$ – $^{13}\text{C}$ – $^{15}\text{N}$ ) G $\alpha$ i3 (Fig. 2*B* and *SI Appendix*, Fig. S1), indicating compound binding. In contrast, another IGGi compound, IGGi-41, that was a potent inhibitor of MDA-MB-231 cell migration (Fig. 1*E*) but did not disrupt GIV–G $\alpha$ i binding (Fig. 2*A*), did not cause NMR signal perturbations (*SI Appendix*, Fig. S2). These results suggested that IGGi-11 binds specifically to G $\alpha$ i3. When IGGi-11-induced NMR perturbations were overlaid on a structural model of IGGi-11-bound G $\alpha$ i3 and compared to a structural model of the GIV–G $\alpha$ i3 complex, several of the amino acids with the largest perturbations (S252, W258, F259, F215, E216, G217, and K35) clustered around the predicted docking site for IGGi-11 and overlapped with the binding area for GIV (Fig. 2*C*). To directly test whether IGGi-11 binds on this predicted site located in the groove between the  $\alpha$ 3 helix and the conformationally dynamic Switch II (SwII) region, we carried out ITC experiments with WT G $\alpha$ i3 or mutants. We found that three different mutations in the predicted binding site for IGGi-11 on G $\alpha$ i3 (F215A, N256E, and W258A) lead to large decreases in compound binding affinity (>10 to 30-fold), whereas another mutation in an amino acid adjacent to the predicted binding site (G42R) did not (Fig. 2*D*). All mutant proteins fold properly and remain functional based on multiple assays (26). The estimated equilibrium dissociation constant ( $K_D$ ) for the G $\alpha$ i3/IGGi-11 interaction based on ITC was  $\sim$ 4  $\mu\text{M}$  (Fig. 2*D*), which was in good agreement with estimates based on curve fits of CSPs observed in NMR experiments (0.9 to 4.6  $\mu\text{M}$ , *SI Appendix*, Fig. S1). IGGi-11 also blocked GIV binding to G $\alpha$ i3 in FP assays with an inhibition constant ( $K_i$ ) of  $\sim$ 14  $\mu\text{M}$  (*SI Appendix*, Fig. S3*A*). Consistently, IGGi-11 also inhibited the ability of GIV to promote the steady-state GTPase activity of G $\alpha$ i3, which reports increased nucleotide exchange in vitro (25) (*SI Appendix*, Fig. S3*B*). Although GIV does not bind to G $\alpha$ o subunits, which belong to the same G $\alpha$ <sub>16</sub> family as G $\alpha$ i3, it does bind to the other G $\alpha$ i isoforms: G $\alpha$ i1 and G $\alpha$ i2 (15, 25). Unsurprisingly, IGGi-11 blocked GIV binding to G $\alpha$ i1 or G $\alpha$ i2 in FP assays with a potency similar to that observed for G $\alpha$ i3 (*SI Appendix*, Fig. S3*A*). Together, these results indicate that IGGi-11 binds to the GIV-interacting site of G $\alpha$ i proteins with low micromolar affinity, thereby precluding the formation of the GIV–G $\alpha$ i complex in vitro.

**IGGi-11 Does Not Affect GIV-Independent Aspects of G-Protein Regulation and Function.** A concern with targeting G $\alpha$ i is the potential on-target but nonetheless undesired effects that may

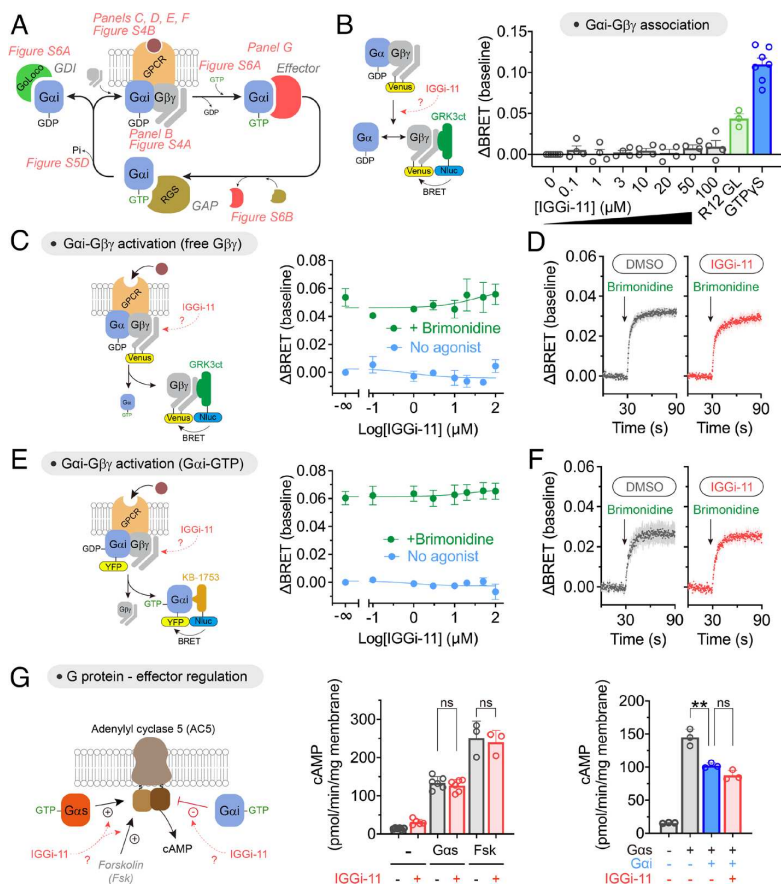
result due to the many functions of G-proteins. The activity of G $\alpha$  subunits depends on the ability to handle nucleotides (GDP/GTP exchange, GTP hydrolysis), on proteins that regulate their activity [G $\beta$  $\gamma$ , GPCRs, guanine nucleotide dissociation inhibitors (GDIs), and GTPase accelerating proteins (GAPs)], or on how they regulate other proteins that propagate signaling (effectors) (Fig. 3*A*). With this in mind, we set out to thoroughly address the potential effect of IGGi-11 on G-protein functions other than those mediated via GIV binding by using isolated cell membranes or purified proteins. First, we tested the effect of IGGi-11 on the association of G $\beta$  $\gamma$  with G $\alpha$  using a bioluminescence resonance energy transfer (BRET) assay (30, 31). We found that concentrations of IGGi-11 up to 100  $\mu\text{M}$  did not cause the dissociation of G $\beta$  $\gamma$  from G $\alpha$ i3 (Fig. 3*B*), whereas incubation with a positive control peptide (R12 GL, 25  $\mu\text{M}$ ) or a nonhydrolyzable GTP analog (GTP $\gamma$ S, 300  $\mu\text{M}$ ) did. Similar observations were made with three other G $\alpha$  subunits that belong to the same family as G $\alpha$ i (i.e., G $\alpha$ o), or to different ones (i.e., G $\alpha$ q and G $\alpha$ 13) (*SI Appendix*, Fig. S4*A*), indicating that IGGi-11 does not disrupt G $\alpha$  $\beta$  $\gamma$  heterotrimers. Using the same assay, we assessed the effect of IGGi-11 on GPCR-mediated activation of G-proteins, which results in the dissociation of G $\beta$  $\gamma$  from G $\alpha$ . We found that concentrations of IGGi-11 up to 100  $\mu\text{M}$  did not interfere with the ability of agonist-stimulated GPCRs to activate G $\alpha$ <sub>13</sub>, G $\alpha$ <sub>o</sub>, G $\alpha$ <sub>q</sub>, or G $\alpha$ <sub>13</sub> heterotrimers (Fig. 3*C* and *SI Appendix*, Fig. S4*B*). Rapid kinetic assays further confirmed that IGGi-11 did not alter the rate of G $\beta$  $\gamma$  dissociation upon GPCR activation (Fig. 3*D*). Moreover, the rate and extent of G $\beta$  $\gamma$ –G $\alpha$ i3 reassociation upon GPCR signal termination was unaffected by IGGi-11 (*SI Appendix*, Fig. S4*C*). Similar observations were made for the reassociation of G $\beta$  $\gamma$  with other G $\alpha$  subunits, like G $\alpha$ o or G $\alpha$ q (*SI Appendix*, Fig. S4*D* and *E*). These observations indicate that IGGi-11 not only fails to disrupt preformed G $\alpha$  $\beta$  $\gamma$  (as in Fig. 3*B*), but also that it does not interfere with the association of trimers. As an alternative to assess GPCR-mediated activation of G-proteins, we used another BRET-based biosensor (32) that directly monitors the formation of GTP-bound G $\alpha$ i3 (Fig. 3*E*). We found that neither amplitude nor kinetics of G $\alpha$ i3–GTP formation upon GPCR stimulation were affected by IGGi-11 (Fig. 3*E* and *F*). We also found that IGGi-11 did not interfere with the spontaneous exchange of GDP for GTP on G $\alpha$ i3 using three independent assays: BRET-based GTP $\gamma$ S binding to G $\alpha$ i in isolated membranes (*SI Appendix*, Fig. S5*A*), binding of fluorescently labeled GTP $\gamma$ S to purified G $\alpha$ i (*SI Appendix*, Fig. S5*B*), or steady-state GTPase activity of purified G $\alpha$ i with radiolabeled GTP (*SI Appendix*, Fig. S5*C*). We also found that IGGi-11 did not affect the hydrolysis of GTP to GDP by purified G $\alpha$ i (*SI Appendix*, Fig. S5*D*).

Next, we evaluated the potential impact of IGGi-11 on the ability of active, GTP-bound G $\alpha$ i proteins to engage and modulate effectors. First, we observed that IGGi-11 did not cause NMR signal perturbations in the  $\alpha$ 3/SwII region of GTP $\gamma$ S-loaded G $\alpha$ i3 (*SI Appendix*, Fig. S6*A*), which contrasts with the observations obtained for GDP-loaded G $\alpha$ i3 (Fig. 2*B* and *SI Appendix*, Fig. S1) and suggests lack of compound binding to active G-proteins. Consistent with this, we also found that IGGi-11 did not inhibit the interaction between purified G $\alpha$ i3 and KB-1753, an effector-like peptide that binds to the  $\alpha$ 3/SwII region of G $\alpha$ i-GTP (33) (*SI Appendix*, Fig. S6*B*). We then tested whether IGGi-11 affected the regulation of a bona fide effector of G $\alpha$ i, i.e., adenylyl cyclase (Fig. 3*G*). In membranes from cells expressing adenylyl cyclase 5, IGGi-11 did not affect either activation mediated by purified G $\alpha$ s or inhibition mediated by purified G $\alpha$ i (Fig. 3*G*). The compound did not affect adenylyl cyclase activity either under basal conditions or upon direct, G-protein-independent activation with forskolin (Fig. 3*G*).



**Fig. 2.** IGGi-11 binding to the GIV-interacting region of Gαi3. (A) IGGi-11 disrupts GIV-Gαi3 binding in pull-down assays. His-Gαi3 was incubated with glutathione agarose-bound GST-GIV (aa 1671-1755) in the presence of the indicated compounds or the positive control NF023 at a concentration of 100  $\mu\text{M}$ . After incubation and washes, bead-bound proteins were separated by SDS-PAGE and immunoblotted (IB) as indicated. Representative of 3 independent experiments. (B) Overlay of  $^1\text{H}$ - $^{15}\text{N}$  TROSY spectra of  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ -Gαi3-GDP in the absence or presence of IGGi-11 are shown on the right. The scatter plot (bottom) corresponds to the quantification of IGGi-11-induced chemical shift perturbations (CSPs). Red, CSP > 5 times the median (M); yellow, CSP > 3xM; blue, CSP < 3xM; gray, no data. Reductions in signal intensity ( $I_{\text{bound}}$ ) below three times the noise (N) are indicated in orange. (C) Comparison of models of IGGi-11 docked onto Gαi3 (Middle and Right, color coded according to NMR perturbations quantified in A) and GIV-bound Gαi3 (Left). (D) Quantification of IGGi-11 binding affinity ( $K_D$ ) for Gαi3 wild type (WT) or the indicated mutants using isothermal titration calorimetry (ITC). Data are representative of at least two independent experiments.





**Fig. 3.** Lack of effect of IGGi-11 on G-protein coupling to GPCRs and effectors. (A) Diagram of key steps and protein interactions involved in Gαi-subunit functions. (B) IGGi-11 does not dissociate Gβγ from Gαi3 in membranes isolated from HEK293T cells expressing a BRET-based biosensor for free Gβγ, whereas two positive controls do (a GoLoco peptide derived from RGS12, R12 GL, 25 μM; and GTPγS 300 μM). (C–F) IGGi-11 does not affect GPCR-mediated activation of Gαi3 as determined by the dissociation of Gαi3-Gβγ heterotrimer (C and D) or the formation of Gαi3-GTP (E and F) using BRET-based biosensors. In C and E, membranes isolated from HEK293T cells expressing the α2A adrenergic receptor were treated with the indicated concentrations of IGGi-11 with (green) or without (blue) stimulation with a receptor agonist (brimonidine, 1 μM) for 2 min before BRET measurements. In D and F, BRET was continuously measured in real time in the presence of 100 μM IGGi-11 or vehicle (1% dimethyl sulfoxide (DMSO), v/v). (G) IGGi-11 does not interfere with G-protein-mediated regulation of adenylyl cyclase. Membranes isolated from HEK293T cells expressing adenylyl cyclase 5 were treated with IGGi-11 (100 μM), purified Gas (0.5 μM), purified myristoylated Gαi1 (Gαi1, 1 μM), and forskolin (Fsk, 10 μM) in the combinations indicated in the graphs. Mean ± SEM (N ≥ 3). \*\*P < 0.01, ANOVA.

Finally, we assessed whether IGGi-11 would preclude the binding of Gαi to other G-protein regulators like GDIs that contain a GoLoco motif (4, 5), or GTPase-accelerating proteins (GAPs) of the regulators of G-protein signaling (RGS) family (6, 7). We found that IGGi-11 did not inhibit the interaction of Gαi3 with the GoLoco motif responsible for the GDI activity of RGS12 (R12 GL, *SI Appendix, Fig. S6B*) or with the GAP RGS4 (*SI Appendix, Fig. S6C*).

Taken together, our results indicate that IGGi-11 specifically inhibits GIV binding to Gαi without interfering with any other major function of Gαi, including nucleotide binding and hydrolysis, association with Gβγ subunits and other cytoplasmic regulators, activation by GPCRs, or modulation of effectors.

#### Validation of an IGGi-11 Analog with Increased Activity in Cells.

After establishing the specificity of IGGi-11 for the target GIV-Gαi complex in vitro, we sought to determine its biological activity in cells. We found that preincubation of MDA-MB-231 cells with IGGi-11 inhibited their ability to migrate only marginally (*SI Appendix, Fig. S7A*). We reasoned that this could be due to low membrane permeability because IGGi-11 contains two

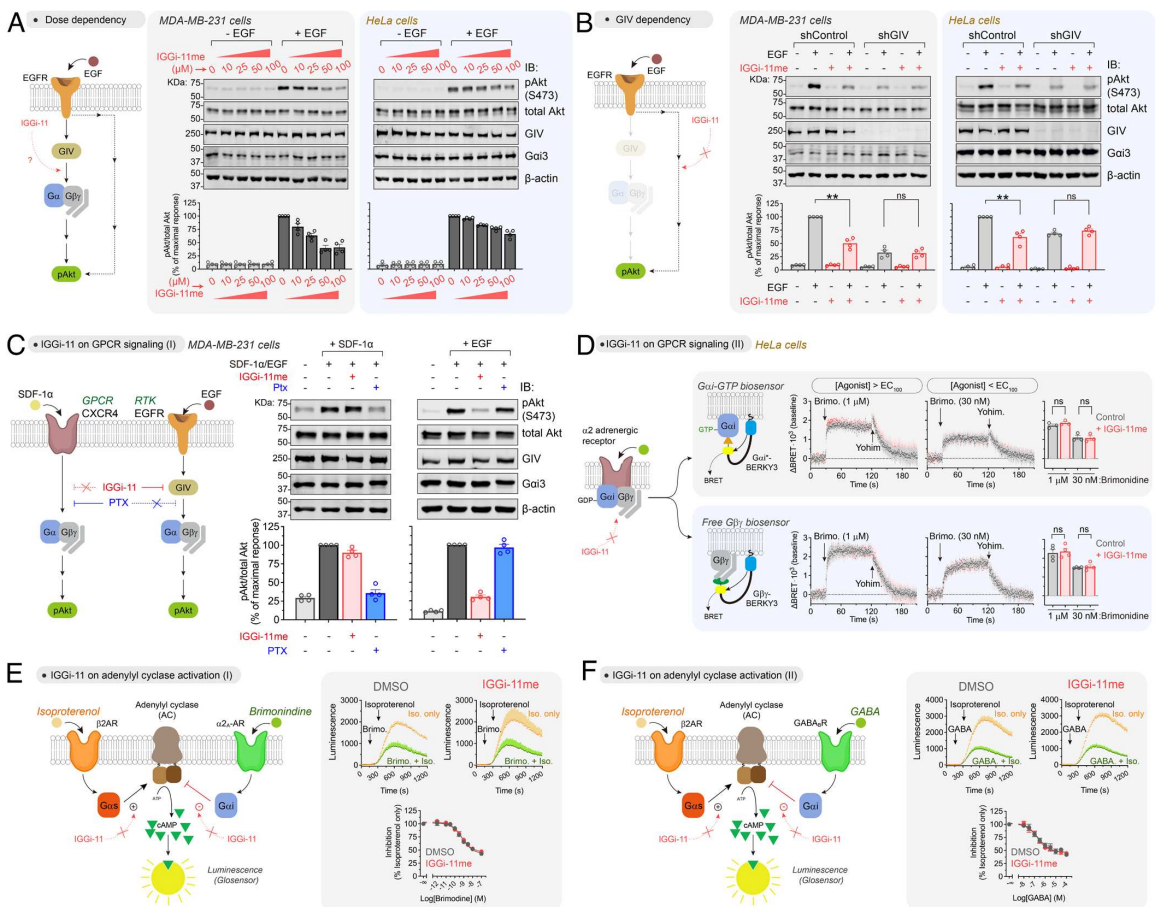
negatively charged carboxylate groups (*SI Appendix, Fig. S7A*). To overcome this, we generated IGGi-11me, an analog in which the carboxylates were esterified with methyl groups. We hypothesized that esterification would increase membrane permeability by eliminating the charges of the carboxylates, and that cytoplasmic esterases would revert the modification to produce IGGi-11, thereby enabling enhanced inhibitory activity in cells (*SI Appendix, Fig. S7A*). Indeed, preincubation of MDA-MB-231 cells with IGGi-11me efficiently reduced their ability to migrate when fetal bovine serum (FBS) was used as the chemoattractant (*SI Appendix, Fig. S7A*). In contrast, IGGi-11me did not affect MDA-MB-231 cell migration when the chemoattractant was SDF-1α, an agonist specific for the G<sub>i</sub>-coupled GPCR CXCR4 (*SI Appendix, Fig. S7A*). This observation is significant for two reasons. First, the effect of IGGi-11me does not blunt MDA-MB-231 cell migration nonspecifically even at a relatively high concentration (up to 100 μM). Second, IGGi-11me does not interfere with mechanisms of migration shared between canonical GPCR-G<sub>i</sub> signaling and signaling mediated by GIV-G<sub>i</sub>, the intended target of IGGi-11me. Also as desired, IGGi-11me (or IGGi-11) did not affect the viability of MDA-MB-231 or MCF-10A cells

(SI Appendix, Fig. S7B), consistent with previous observations that depletion of GIV does not affect MDA-MB-231 cell growth under similar conditions (17). We confirmed that IGGi-11me had higher permeability than IGGi-11 by using a parallel artificial membrane permeability assay (SI Appendix, Fig. S7C). We also confirmed that IGGi-11me was converted to IGGi-11 by esterases present in the cytosol of MDA-MB-231 cells (SI Appendix, Fig. S7D), which is a critical step because IGGi-11me would be a poor inhibitor of GIV–Gα3 binding at the concentrations tested in cells due to its lower potency compared to IGGi-11 (SI Appendix, Fig. S7E). These results indicate that IGGi-11me serves as a prodrug that allows the action of the active GIV–Gα3 inhibitor compound, IGGi-11, in cells.

### IGGi-11me Inhibits GIV-Dependent Cancer Cell Signaling.

Previous work has shown that GIV mediates the activation of Akt downstream of various receptor tyrosine kinases (RTKs),

including the epidermal growth factor receptor (EGFR), and other surface receptors via G-protein (i.e., Gβγ)-dependent activation of PI3K (8, 15, 17, 23, 25, 34–36). We found that IGGi-11me reduced, in a dose-dependent manner, the phosphorylation of Akt at S473 (pAkt) upon epidermal growth factor (EGF) stimulation in two cell lines, MDA-MB-231 and HeLa, indicating reduced Akt activity (Fig. 4A). The lack of complete Akt inhibition is consistent with the known existence of GIV-independent mechanisms utilized by EGFR to activate PI3K–Akt signaling (37). In fact, the extent of IGGi-11me-mediated inhibition of Akt was similar to that observed upon depletion of GIV in these cell lines (Fig. 4B). Moreover, IGGi-11me failed to further reduce Akt activation in GIV-depleted cells even at the maximal concentration of compound tested (100 μM), indicating that it does not affect GIV-independent mechanisms of Akt activation downstream of EGFR (Fig. 4B). These results suggest that, even at relatively high



**Fig. 4.** IGGi-11me specifically inhibits GIV-dependent G-protein cell signaling. (A) IGGi-11me inhibits EGF-stimulated Akt activation (phospho-serine 473, pAkt S473) in MDA-MB-231 and HeLa cells. Cells were preincubated with the indicated concentrations of IGGi-11me and stimulated with EGF (1.6 nM for MDA-MB-231 or 50 nM for HeLa) for 5 min before lysis and immunoblotting. (B) IGGi-11me (100 μM) does not inhibit EGF-stimulated Akt activation in GIV-depleted cells. GIV-depleted cells (shGIV) or control cells (shControl) were treated as in A. (C) IGGi-11me does not block Akt activation upon stimulation of the GPCR CXCR4. MDA-MB-231 cells were preincubated with IGGi-11me (100 μM) or pertussis toxin (PTX, 100 ng/mL) and stimulated with SDF-1α (100 ng/mL for 10 min) or EGF (1.6 nM for 5 min) before processing as in A. (D) IGGi-11me does not affect GPCR-mediated modulation of G-protein activity. HeLa cells expressing BRET biosensors for Gαi-GTP (Gαi<sup>B</sup>-BERKY3) or free Gβγ (Gβγ-BERKY3) were preincubated with IGGi-11me (100 μM) and sequentially treated with the α<sub>2</sub> adrenergic agonist brimonidine and the antagonist yohimbine (25 μM) during real-time BRET measurements as indicated in the figure. (E and F) IGGi-11me does not affect GPCR-mediated modulation of cAMP responses. HEK293T cells expressing Glosensor, a luminescence-based cAMP sensor, and either α<sub>2A</sub>-AR (E) or GABA<sub>A</sub>R (F) were pretreated with IGGi-11me (100 μM) or DMSO before measuring luminescence. Cells were treated with isoproterenol (100 nM) with or without prestimulation with brimonidine (E) or GABA (F) as indicated. The concentration of brimonidine and GABA in the kinetic traces shown was 1 μM and 100 μM, respectively. All results are mean ± SEM (N ≥ 3). \*\*P < 0.01; ns, P > 0.05, ANOVA.

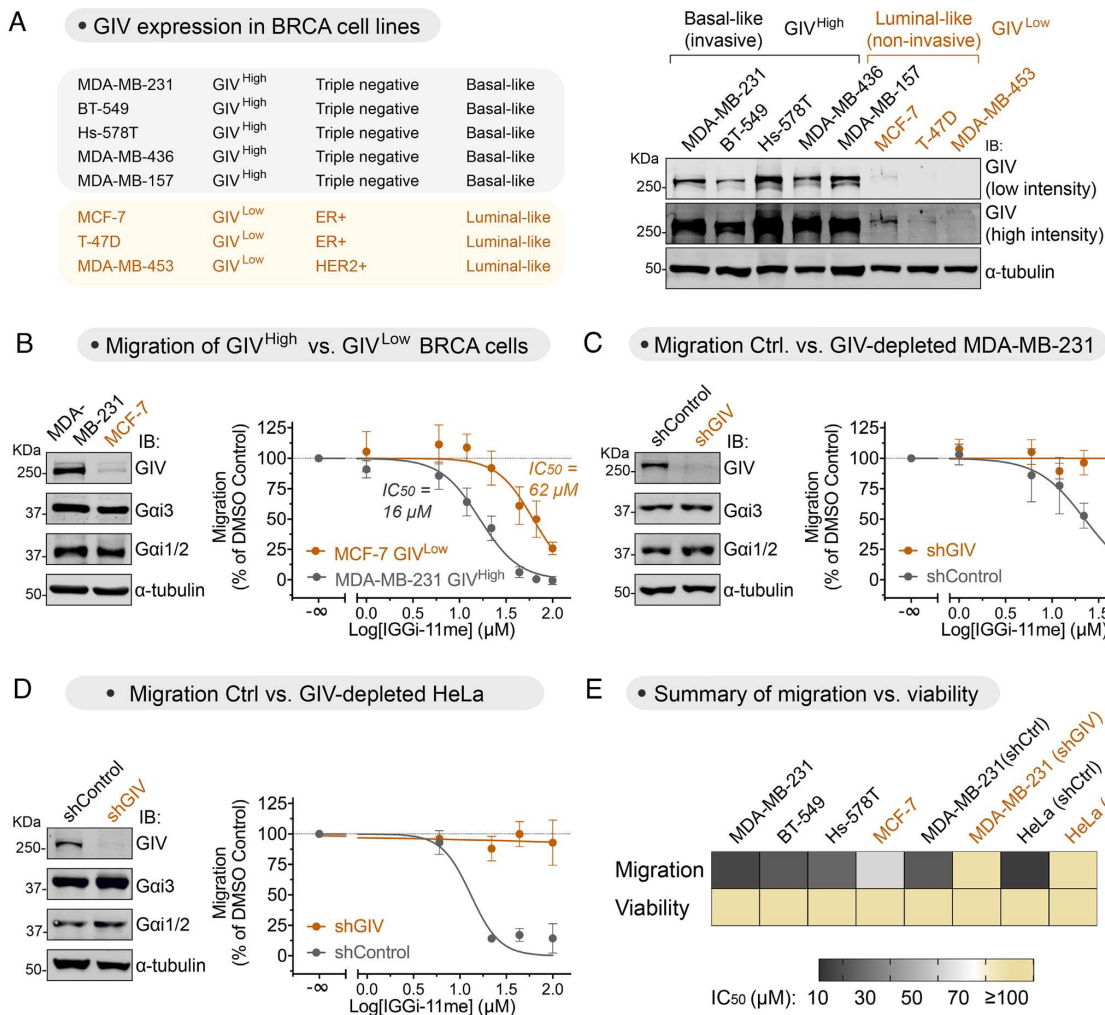
concentrations, IGGi-11me does not have nonspecific effects on Akt signaling because it lacks an inhibitory effect in the absence of the intended target. Also, IGGi-11me did not change the total amount of GIV or G $\alpha$ i (Fig. 4 A and B), supporting that its mechanism of action is the disruption of the interaction of the two proteins, rather than indirectly altering their abundance. These results are consistent with the idea that IGGi-11me specifically inhibits GIV-dependent G-protein signaling in cancer cells.

**IGGi-11me Does Not Affect GIV-Independent G-Protein Cell Signaling.** Next, we set out to further assess the specificity of IGGi-11me in cell signaling. Although IGGi-11 does not interfere with GIV-independent mechanisms of G-protein regulation in vitro (Fig. 3 and *SI Appendix, Figs. S4–S6*), confirmation that the same holds for IGGi-11me in cells was warranted given the relatively high concentrations of compounds needed to block GIV-dependent signaling. First, we compared side by side the effect of a maximal concentration of IGGi-11me (100  $\mu$ M) on GIV-dependent and GIV-independent G-protein signaling in the same cell line (MDA-MB-231) with the same readout (Akt activation). For GIV-dependent G-protein signaling, we stimulated cells with EGF as in Fig. 4 A and B, whereas for GIV-independent G-protein signaling, we stimulated cells with SDF-1 $\alpha$ , an agonist for the endogenously expressed G $_i$ -coupled GPCR CXCR4 (Fig. 4C). We found that IGGi-11me inhibited Akt activation in response to EGF but not in response to SDF-1 $\alpha$  (Fig. 4C), indicating that it does not interfere with GPCR-mediated G-protein signaling. In contrast, PTX, which precludes G $\alpha$ i activation by GPCRs but not by GIV (38), efficiently blocked activation of Akt in response to SDF-1 $\alpha$  but not to EGF (Fig. 4C). These results indicate that IGGi-11me specifically targets GIV-dependent G-protein signaling mechanisms in cells without interfering with canonical GPCR-mediated G-protein signaling. This result is in good agreement with the lack of effect of IGGi-11me on SDF-1 $\alpha$ -stimulated migration of MDA-MB-231 cell (*SI Appendix, Fig. S7A*), further supporting the notion that the compound does not affect signaling mechanisms shared between GPCR-G $_i$  and GIV-G $_i$  pathways. To substantiate this point, we assessed the effect of IGGi-11me on GPCR signaling by using BRET-based biosensors that directly monitor the activation of endogenous G-proteins (32). More specifically, HeLa cells expressing biosensors for either G $\alpha$ i-GTP or free G $\beta\gamma$  were treated with a maximal concentration of IGGi-11me (100  $\mu$ M) exactly under the same conditions that led to decreased GIV-dependent Akt activation after EGF stimulation in this cell line (Fig. 4 A and B). We found that G-protein responses elicited by stimulation of endogenous  $\alpha$ 2 adrenergic receptors with maximal (>EC $_{100}$ ) or submaximal (<EC $_{100}$ ) concentrations of a cognate agonist were unaltered by IGGi-11me (Fig. 4D). Not only were the amplitudes and rates of the activation responses unchanged, but the rates of deactivation upon GPCR blockade with an antagonist also remained the same (Fig. 4D), indicating that the compound does not have effects on nucleotide exchange or hydrolysis rates, or the dissociation or reassociation of G-protein heterotrimers upon GPCR-mediated modulation. We went on to determine whether, in addition to not having an effect at the level of G-protein regulation, IGGi-11me also lacked an effect on a well-established downstream signaling readout like the second messenger cAMP. For this, we measured cAMP regulation by GPCRs in HEK293T cells in the presence and absence of a maximal concentration of IGGi-11me (100  $\mu$ M). Endogenous  $\beta$ -adrenergic receptors were stimulated with isoproterenol to elevate cAMP levels via G $\alpha$ s, and this response was modulated by the stimulation of exogenously expressed  $\alpha$ 2 $_A$  adrenergic receptors (Fig. 4E) or GABA $_B$  receptors

(Fig. 4F), which suppress the isoproterenol-stimulated cAMP response via G $\alpha$ i. We found that IGGi-11me had no effect on 1) the cAMP response to isoproterenol in the absence of G $_i$ -mediated regulation, or 2) the efficacy or potency of G $_i$ -mediated inhibition of the isoproterenol response by any of the two GPCRs tested (Fig. 4 E and F). Taken together, these results show that IGGi-11me does not interfere with GIV-independent G-protein signaling, including that elicited by canonical GPCR/G-protein signaling pathways.

**IGGi-11me Specifically Inhibits GIV-Dependent Tumor Cell Migration.** Previous evidence indicates that GIV is expressed at high levels in metastatic cancers, and that formation of the GIV–G $\alpha$ i complex favors cell migration (15, 18–20). Consistent with some of these observations, we found that invasive breast cancer (BRCA) cell lines prone to metastasis expressed higher levels of GIV (GIV $^{\text{High}}$ ) than noninvasive breast cancer cell lines (GIV $^{\text{Low}}$ ) (Fig. 5A). IGGi-11me was approximately four times more potent in inhibiting the migration of MDA-MB-231 cells (GIV $^{\text{High}}$ ) than that of MCF-7 cells (GIV $^{\text{Low}}$ ) (Fig. 5B). This difference in IGGi-11me sensitivity could not be attributed to differences in G $\alpha$ i protein abundance because they were present in similar amounts in both cell lines (Fig. 5B). While we could not test the effect of IGGi-11me on the GIV $^{\text{Low}}$  cell lines T47D and MDA-MB-453 because they lacked measurable migration, we found that IGGi-11me inhibited cell migration in the GIV $^{\text{High}}$  cell lines BT-549 and Hs578T with a potency similar to that seen for MDA-MB-231 cells (*SI Appendix, Fig. S8A*). Thus, despite the different genetic background of these cell lines bearing different drivers of cancer traits, the common denominator is that GIV expression (and presumably the formation of a GIV–G $\alpha$ i complex) correlates with sensitivity to IGGi-11me. To further assess the specificity of IGGi-11me in inhibiting GIV-dependent tumor cell migration, we tested its effect on GIV-depleted MDA-MB-231 cells. We found that, compared to control cells, IGGi-11me had no effect on MDA-MB-231 cell migration upon GIV depletion even when tested at maximal concentration (100  $\mu$ M) (Fig. 5C). Similar observations were made with GIV-depleted HeLa cells, which are of different origin and genetic background than that of MDA-MB-231 (Fig. 5D). Moreover, GIV-depleted MDA-MB-231 cells or HeLa cells contained amounts of G $\alpha$ i proteins similar to those in their corresponding control cells (Fig. 5 C and D), further supporting that the inhibition of cell migration exerted by IGGi-11me is GIV-dependent. These results are also consistent with the findings shown in *SI Appendix, Fig. S7A* indicating that IGGi-11me does not have an effect on CXCR4-promoted cell migration, which rule out that IGGi-11me affects mechanisms of migration shared between canonical GPCR-G $_i$  signaling and signaling mediated by GIV-G $_i$ . Furthermore, the inhibition of cell migration by IGGi-11me was not a consequence of reduced cell viability, as the latter was not affected by the compound in any of the cell lines investigated (Fig. 5E and *SI Appendix, Fig. S8B*). These findings indicate that IGGi-11me specifically blocks GIV-dependent tumor cell migration, implying that the disruption of the GIV–G $\alpha$ i complex hinders the proinvasive features of GIV $^{\text{High}}$  cancer cells.

**IGGi-11me Inhibits Cancer Cell Growth in Tumor-Like Contexts.** GIV-depleted MDA-MB-231 cells fail to metastasize in mouse xenograft models, which correlates well with the effects of GIV depletion on tumor cell migration and invasion (21). Unfortunately, we could not test the effect of IGGi-11me on cancer mouse models to assess metastasis because preliminary results showed that IGGi-11me is rapidly degraded in plasma. As an alternative, we set out to investigate the impact of IGGi-11me



**Fig. 5.** IGGI-11me blocks GIV-dependent tumor cell migration. (A) Basal-like invasive breast cancer (BRCA) cell lines express higher amounts of GIV (GIV<sup>High</sup>) than luminal-like noninvasive BRCA cell lines (GIV<sup>Low</sup>) as determined by immunoblotting. (B) IGGI-11me inhibits cell migration more potently in MDA-MB-231 cells (GIV<sup>High</sup>) than in MCF-7 cells (GIV<sup>Low</sup>). Chemotactic cell migration toward FBS was determined in the presence of the indicated concentrations of IGGI-11me using a modified Boyden-chamber assay. (C and D) IGGI-11me-mediated inhibition of tumor cell migration is lost upon depletion of GIV from MDA-MB-231 (C) or HeLa (D) cells. GIV-depleted cells (shGIV) or control cells (shControl) were processed as described in B. (E) IGGI-11me impairs tumor cell migration without affecting cell viability. Heatmap comparing the half-maximal inhibitory concentration (IC<sub>50</sub>) of IGGI-11me on cell migration or viability of the indicated cell lines. IC<sub>50</sub> values were determined from results shown in this figure or in *SI Appendix, Fig. S8*. Cell viability was determined upon incubation with IGGI-11me for 24 h, which is longer than the times cells were exposed to the compound in cell migration assays. All results are mean ± SEM (N ≥ 3).

on the behavior of MDA-MB-231 cells in a tumor-like context. This was motivated by previous observations that loss of GIV does not affect the growth of tumor cells, including MDA-MB-231, on plastic dishes (17, 21), but hinders growth in three-dimensional Matrigel cultures (17), which account for tumor cell interactions with the extracellular matrix and recapitulate many of the behavioral features of cancer cells in tumors in situ (39). We found that IGGI-11me mimicked previous observations (17) upon loss of GIV in Matrigel cultures—i.e., MDA-MB-231 became smaller and more organized acinar structures than control cells, resulting in an overall reduction of cell growth (*SI Appendix, Fig. S9A and B*). In contrast, IGGI-11me did not affect the growth of nontransformed MCF-10A breast cells in Matrigel cultures even when tested at a maximal concentration (100 μM) (*SI Appendix, Fig. S9B*), suggesting that the effect on

MDA-MB-231 was not due to nonspecific toxicity. As a second model to assess the effect of IGGI-11me on cancer cell growth in a tumor-like environment, we pretreated MDA-MB-231 cells with IGGI-11me and assessed their ability to form tumors when implanted subcutaneously as xenografts in mice (*SI Appendix, Fig. S9C*). We found that IGGI-11me-treated cells formed tumors less efficiently than controls (*SI Appendix, Fig. S9C*). In contrast, when cells treated and prepared in the same way as above were seeded on plastic dishes and grown under standard cell culture conditions, IGGI-11me had no effect (*SI Appendix, Fig. S9C*), suggesting that the effect on xenograft growth is specific to the tumor-like context for MDA-MB-231 cells recapitulated in mice. Together, these results indicate that disruption of the GIV–Gai interaction by IGGI-11me prevents cancer cell growth in tumor-like contexts.

## Discussion

In this work, we identify and characterize a chemical probe of broad utility for dissecting atypical mechanisms of cellular communication mediated by G-proteins with important biomedical implications not only for cancer, but also for fibrosis, and male fertility, among other maladies (8, 9, 17, 18, 20, 40, 41). From a broader perspective, this work provides the proof of principle for a modality of pharmacological targeting in heterotrimeric G-protein signaling that deviates from the widespread focus on GPCRs or the direct ablation of G-protein activity *en toto*. This modality consists of targeting G-proteins to selectively disrupt specific mechanisms by which they are regulated. IGGi-11 disrupts G $\alpha$ i binding to GIV but not to many of its other binding partners, despite them physically engaging the same region of G $\alpha$ i as GIV. This region includes the SwII, which is dynamic and adopts different conformations depending on the protein partner bound to G $\alpha$ i. Although it is tempting to speculate that the selectivity of IGGi-11 may arise from its relative ability to interact with these different conformations, the structural basis for the action of IGGi-11 remains to be fully elucidated. The targeting modality described here follows the path opened by recent advances on small-molecule inhibitors for another GTPase, KRas, in reshaping the traditional definition of what constitutes a druggable target (42, 43). The main limitations of IGGi-11(me) relate to its modest affinity and poor stability in plasma. Because its chemical scaffold is synthetically tractable, IGGi-11 may further serve as a lead compound to develop analogs with improved potency and pharmacokinetic properties that could have therapeutic value.

## Materials and Methods

Chemical compounds of interest were purchased from reliable vendors or synthesized in-house, and tested in *in vitro* assays, including NMR, BRET assays, or different protein-protein binding experiments following previously established

procedures that are described in detail in *SI Appendix*. Cell-based experiments to assess the efficacy and specificity of compounds were also carried out using previously established procedures and/or cell lines, including cell migration assays using modified Boyden chambers, immunoblotting, and signaling assays, all of which are described in detail in *SI Appendix* along with the animal experiments measuring xenograft tumor growth by luminescence bioimaging.

**Data, Materials, and Software Availability.** All study data and protocols are included and can be accessed directly in the article and/or *SI Appendix*. No code or software was generated for this study, and data were not deposited in a public database.

**ACKNOWLEDGMENTS.** This work was supported by NIH grant R01GM130120 and the Karin Grunebaum Cancer Research Foundation (to M.G.-M). J.Z. was supported by a Dahod International Scholar Award, and A.L. was supported by a F31 Ruth L. Kirschstein NRSA Predoctoral Fellowship (F31NS115318). F.J.B. was supported by Spanish Government grant PID2020-113225GB-I00. M.F.-G. was supported by Spanish Government fellowship PRE-2018-085788. We thank the ICCB-Longwood Screening Facility at Harvard Medical School, S. Whelan (Boston University), F. Seta (Boston University), N. Ganem (Boston University), and J. B. Blanco-Canosa (Institute for Advanced Chemistry of Catalonia) for access to instrumentation and reagents. We thank N. Merino (CIC bioGUNE, Spain) for help with the purification of G $\alpha$ i3 protein used in NMR studies, M. Rico (CIC bioGUNE, Spain) for access to NMR spectrometers, and A. González-Magaña (CIC bioGUNE, Spain) for preliminary ITC experiments. We thank the following investigators for providing DNA plasmids: K. Martemyanov (The Scripps Research Institute, Jupiter, FL), N. Lambert (Augusta University, Augusta, GA), P. Wedegaertner (Thomas Jefferson University, Philadelphia, PA), J. Blumer (Medical University of South Carolina, Charleston, SC), J. Sondek (University of North Carolina, Chapel Hill, NC), N. Artemyev (University of Iowa), C. Dressauer (University of Texas Health Science Center at Houston, TX), and M. Linder (Cornell University).

Author affiliations: <sup>a</sup>Department of Biochemistry & Cell Biology, Boston University, Chobanian & Avedisian School of Medicine, Boston, MA 02118; <sup>b</sup>Centro de Investigaciones Biológicas-Centro Superior de Investigaciones Científicas, Madrid, Spain; <sup>c</sup>Department of Chemistry, Boston University, College of Arts & Sciences, Boston, MA 02115; <sup>d</sup>German Center for Neurodegenerative Diseases, Göttingen, Germany; and <sup>e</sup>Department of Biology, College of Arts & Sciences, Boston University, Boston, MA 02115

1. W. I. Weis, B. K. Koblick, The molecular basis of G protein-coupled receptor activation. *Annu. Rev. Biochem.* **87**, 897–919 (2018).
2. A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schiøth, D. E. Gloriam, Trends in GPCR drug discovery: New agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017).
3. A. G. Gilman, G proteins: Transducers of receptor-generated signals. *Annu. Rev. Biochem.* **56**, 615–649 (1987).
4. M. Sato, J. B. Blumer, V. Simon, S. M. Lanier, Accessory proteins for G proteins: Partners in signaling. *Annu. Rev. Pharmacol. Toxicol.* **46**, 151–187 (2006).
5. D. P. Siderovski, F. S. Willard, The GAPs, GEFs, and GDI of heterotrimeric G-protein alpha subunits. *Int. J. Biol. Sci.* **1**, 51–66 (2005).
6. E. M. Ross, T. M. Wilkie, GTPase-activating proteins for heterotrimeric G proteins: Regulators of G protein signaling (RGS) and RGS-like proteins. *Annu. Rev. Biochem.* **69**, 795–827 (2000).
7. H. G. Dohlman, J. Thorne, RGS proteins and signaling by heterotrimeric G proteins. *J. Biol. Chem.* **272**, 3871–3874 (1997).
8. M. Garcia-Marcos, P. Ghosh, M. G. Farquhar, GIV/Girdin transmits signals from multiple receptors by triggering trimeric G protein activation. *J. Biol. Chem.* **290**, 6697–6704 (2015).
9. V. DiGiorgio, A. Marivin, M. Garcia-Marcos, When heterotrimeric G proteins are not activated by G protein-coupled receptors: Structural insights and evolutionary conservation. *Biochemistry* **57**, 255–257 (2018).
10. M. J. Cismowski *et al.*, Activation of heterotrimeric G-protein signaling by a ras-related protein. Implications for signal integration. *J. Biol. Chem.* **275**, 23421–23424 (2000).
11. G. G. Tall, Ric-8 regulation of heterotrimeric G proteins. *J. Recept. Signal Transduct. Res.* **33**, 139–143 (2013).
12. A. P. Campbell, A. V. Smrcka, Targeting G protein-coupled receptor signalling by blocking G proteins. *Nat. Rev. Drug Discov.* **17**, 789–803 (2018).
13. T. M. Bonacci *et al.*, Differential targeting of Gbetagamma-subunit signaling with small molecules. *Science* **312**, 443–446 (2006).
14. E. Kostenis, E. M. Pfeil, S. Annala, Heterotrimeric Gq proteins as therapeutic targets? *J. Biol. Chem.* **295**, 5206–5215 (2020).
15. M. Garcia-Marcos, P. Ghosh, M. G. Farquhar, GIV is a nonreceptor GEF for G alpha i with a unique motif that regulates Akt signaling. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3178–3183 (2009).
16. A. Leyme *et al.*, Specific inhibition of GPCR-independent G protein signaling by a rationally engineered protein. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E10319–E10328 (2017).
17. A. Leyme, A. Marivin, L. Perez-Gutierrez, L. T. Nguyen, M. Garcia-Marcos, Integrins activate trimeric G proteins via the nonreceptor protein GIV/Girdin. *J. Cell Biol.* **210**, 1165–1184 (2015).
18. P. Ghosh, M. Garcia-Marcos, M. G. Farquhar, GIV/Girdin is a rheostat that fine-tunes growth factor signals during tumor progression. *Cell Adh. Migr.* **5**, 237–248 (2011).
19. M. Garcia-Marcos *et al.*, Expression of GIV/Girdin, a metastasis-related protein, predicts patient survival in colon cancer. *FASEB J.* **25**, 590–599 (2011).
20. P. Ghosh, Heterotrimeric G proteins as emerging targets for network-based therapy in cancer: End of a long futile campaign striking heads of a Hydra. *Aging* **7**, 469–474 (2015).
21. P. Jiang *et al.*, An actin-binding protein Girdin regulates the motility of breast cancer cells. *Cancer Res.* **68**, 1310–1318 (2008).
22. A. Leyme, A. Marivin, M. Garcia-Marcos, GIV/Girdin (Galpha-interacting, vesicle-associated Protein/Girdin) creates a positive feedback loop that potentiates outside-in integrin signaling in cancer cells. *J. Biol. Chem.* **291**, 8269–8282 (2016).
23. M. Garcia-Marcos *et al.*, Functional characterization of the guanine nucleotide exchange factor (GEF) motif of GIV protein reveals a threshold effect in signaling. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1961–1966 (2012).
24. G. S. Ma *et al.*, Therapeutic effects of cell-permeant peptides that activate G proteins downstream of growth factors. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E2602–E2610 (2015).
25. M. Garcia-Marcos, P. Ghosh, J. Ear, M. G. Farquhar, A structural determinant that renders G alpha(i) sensitive to activation by GIV/Girdin is required to promote cell migration. *J. Biol. Chem.* **285**, 12765–12777 (2010).
26. A. I. de Opakua *et al.*, Molecular mechanism of Galpha(i) activation by non-GPCR proteins with a Galpha(i)-Binding and Activating motif. *Nat. Commun.* **8**, 15163 (2017).
27. V. DiGiorgio *et al.*, The Galpha(i)-GIV binding interface is a druggable protein-protein interaction. *Sci. Rep.* **7**, 8575 (2017).
28. N. A. Kalogiropoulos *et al.*, Structural basis for GPCR-independent activation of heterotrimeric G proteins. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 16394–16403 (2019).
29. B. Hollins, S. Kuravi, G. J. Digby, N. A. Lambert, The c-terminus of GRK3 indicates rapid dissociation of G protein heterotrimers. *Cell. Signal.* **21**, 1015–1021 (2009).
30. I. Masuho *et al.*, Distinct profiles of functional discrimination among G proteins determine the actions of G protein-coupled receptors. *Sci. Signal.* **8**, ra123 (2015).
31. M. Maziarz *et al.*, Revealing the activity of trimeric G-proteins in live cells with a versatile biosensor design. *Cell* **182**, 770–785 e716 (2020).
32. C. A. Johnston *et al.*, Minimal determinants for binding activated G alpha from the structure of a G alpha(i1)-peptide dimer. *Biochemistry* **45**, 11390–11400 (2006).
33. P. Ghosh *et al.*, A Galpha(i)-GIV molecular complex binds epidermal growth factor receptor and determines whether cells migrate or proliferate. *Mol. Biol. Cell* **21**, 2338–2354 (2010).

34. C. Lin *et al.*, Structural basis for activation of trimeric Gi proteins by multiple growth factor receptors via GIV/Girdin. *Mol. Biol. Cell* **25**, 3654–3671 (2014).
35. M. Garcia-Marcos, J. Ear, M. G. Farquhar, P. Ghosh, A GDI (AGS3) and a GEF (GIV) regulate autophagy by balancing G protein activity and growth factor signals. *Mol. Biol. Cell* **22**, 673–686 (2011).
36. M. A. Lemmon, J. Schlessinger, Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117–1134 (2010).
37. M. Garcia-Marcos, Complementary biosensors reveal different G-protein signaling modes triggered by GPCRs and non-receptor activators. *ELife* **10**, e65620 (2021).
38. J. Debnath, J. S. Brugge, Modelling glandular epithelial cancers in three-dimensional cultures. *Nat. Rev. Cancer* **5**, 675–688 (2005).
39. A. J. Minn *et al.*, Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).
40. S. Reynoso *et al.*, GIV/Girdin, a non-receptor modulator for Galphai3, regulates spatiotemporal signaling during sperm capacitation and is required for male fertility. *ELife* **10**, e69160 (2021).
41. I. Lopez-Sanchez *et al.*, GIV/Girdin is a central hub for profibrogenic signalling networks during liver fibrosis. *Nat. Commun.* **5**, 4451 (2014).
42. J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells, K. M. Shokat, K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).
43. F. Skoulidis *et al.*, Sotorasib for lung cancers with KRAS p. G12C mutation. *N. Engl. J. Med.* **384**, 2371–2381 (2021).



## Structural analysis of ING3 protein and histone H3 binding

Mariola Ferreras-Gutiérrez<sup>a</sup>, Belén Chaves-Arquero<sup>a</sup>, Amaia González-Magaña<sup>b</sup>, Nekane Merino<sup>c</sup>, Ignacio Amusatogui-Mateu<sup>a</sup>, Sonia Huecas<sup>a</sup>, Francisco J. Medrano<sup>a</sup>, Francisco J. Blanco<sup>a,\*</sup>

<sup>a</sup> Centro de Investigaciones Biológicas Margarita Salas (CIB), CSIC, Madrid 28040, Spain

<sup>b</sup> Instituto Biofísica and Departamento de Bioquímica y Biología Molecular (CSIC, UPV/EHU), University of the Basque Country, 48940 Leioa, Spain

<sup>c</sup> CIC bioGUNE, Parque Tecnológico de Bizkaia, 48160 Derio, Spain

### ARTICLE INFO

#### Keywords:

ING3  
Histone H3  
Protein structure

### ABSTRACT

Proteins belonging to the ING family regulate the transcriptional state of chromatin by recruiting remodeling complexes to sites with histone H3 trimethylated at Lysine 4 (H3K4me3). This modification is recognized by the Plant HomeoDomain (PHD) present at the C-terminal region of the five ING proteins. ING3 facilitates acetylation of histones H2A and H4 by the NuA4-Tip60 MYST histone acetyl transferase complex, and it has been proposed to be an oncoprotein. The crystal structure of the N-terminal domain of ING3 shows that it forms homodimers with an antiparallel coiled-coil fold. The crystal structure of the PHD is similar to those of its four homologs. These structures explain the possible deleterious effects of ING3 mutations detected in tumors. The PHD binds histone H3K4me3 with low-micromolar, and binds the non-methylated histone with a 54-fold reduced affinity. Our structure explains the impact of site directed mutagenesis experiments on histone recognition. These structural features could not be confirmed for the full-length protein as solubility was insufficient for structural studies, but the structure of its folded domains suggest a conserved structural organization for the ING proteins as homodimers and bivalent readers of the histone H3K4me3 mark.

### 1. Introduction

In the nucleosome, the fundamental unit of the chromatin, two superhelical turns of DNA are wound around an octamer of the four core histones. The N-terminal regions of the histones are flexible and rich in positive charges, and interact with the negatively charged DNA phosphate backbone increasing the compactness of the chromatin [1]. This compactness is dynamic and largely regulated by post-translational modifications of histones, most frequently at their disordered tails. These modifications may directly affect the level of chromatin compaction or may recruit specific chromatin remodeling complexes, affecting DNA replication, repair, or transcription [2,3]. Histone tails are selectively modified by enzymatic complexes containing proteins with domains that in turn recognize one or more of the possible histone modification types [4].

The Inhibitors of Growth (ING) human gene family consists of five homologs [5], coding for proteins that recruit chromatin remodeling complexes to sites with histone H3 trimethylated at Lysine 4 (H3K4me3). This modification is frequent in the promoter-proximal

region of transcriptionally active genes [6]. ING1 and ING2 form part of histone deacetylase (HDAC) complexes, which remove acetyl groups from lysine side chains, restoring their positive charge, increasing chromatin compactness, and repressing gene transcription. ING3, 4, and 5 form part of histone acetyl transferase (HAT) complexes that acetylate lysine side chains, removing their positive charge, decreasing chromatin compactness, and increasing gene transcription [7]. These enzymatic complexes can also affect gene transcription through acetylation/deacetylation of proteins other than histones, such as p53 [8,9]. *In vitro* and *in vivo* studies suggest that ING1 and ING2 are tumor suppressor genes, ING3 is an oncogene, and ING4 and ING5 do not clearly fall in these two simple categories. These findings prompted a proposal to rename the family as INstructors of Growth [10].

The predominant isoforms of the ING proteins have a conserved N-terminal region, a central non-conserved region containing the Nuclear Localization Signal (NLS), and a conserved C-terminal PHD. The histone H3K4me3 post-translational modification is recognized by the PHD of the ING proteins [11]. Crystal structures of the PHD of ING1 [12], ING2 [13], ING4 [14], and ING5 [15] bound to H3K4me3 fragments display

\* Corresponding author at: CIB-CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain.

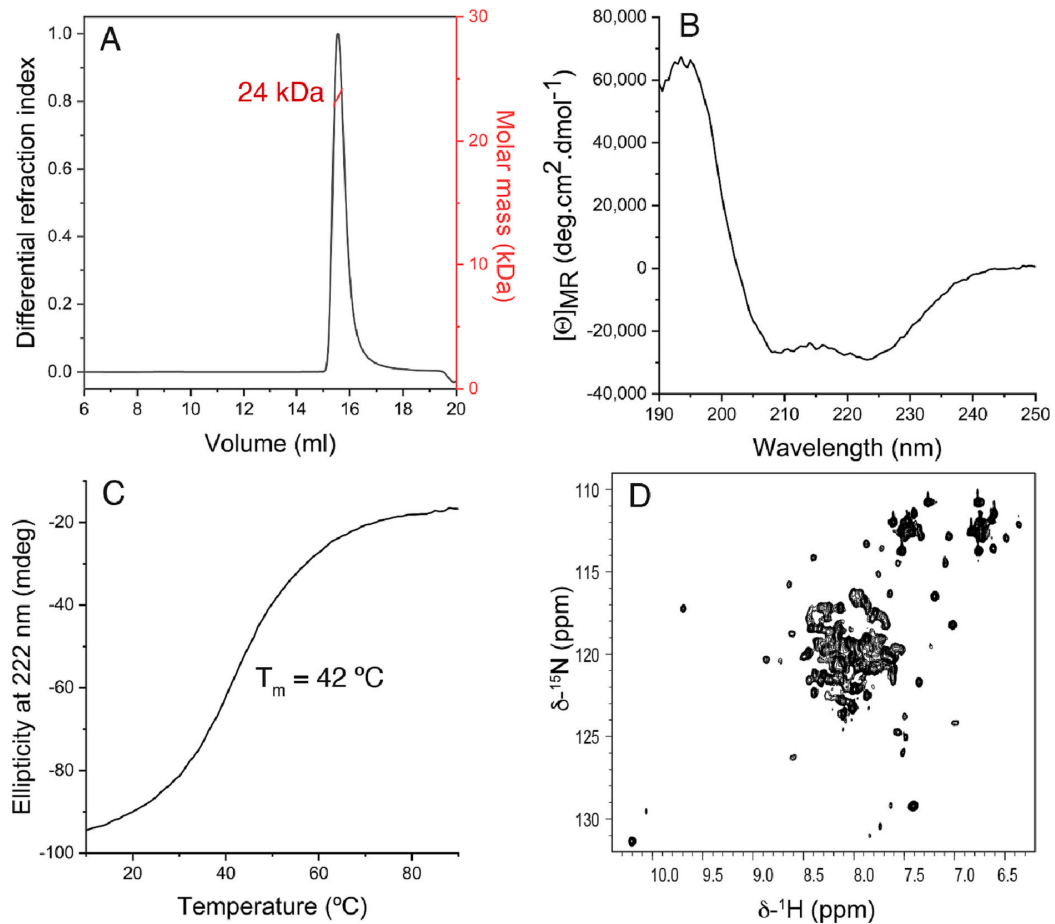
E-mail address: [fj.blanco@cib.csic.es](mailto:fj.blanco@cib.csic.es) (F.J. Blanco).

<https://doi.org/10.1016/j.ijbiomac.2023.124724>

Received 6 March 2023; Received in revised form 21 April 2023; Accepted 30 April 2023

Available online 4 May 2023

0141-8130/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



**Fig. 1.** The Nt domain of ING3 forms dimers with coiled-coil helical structure. (A) SEC-MALS analysis at 23 °C. (B) CD spectrum of ING3 Nt at 25 °C. (C) Thermal denaturation of ING3 Nt. (D)  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectrum of ING3 Nt at 25 °C.

broadly conserved features with minor differences in the conformation of the H3K4me3 peptides [14]. The structural organization of the full-length ING4 and ING5 proteins has been experimentally studied [16,17]. They form homodimers, with the N-terminal domain folded into an antiparallel coiled-coil [18] while the central region, connecting the two folded domains, is disordered and binds double-stranded DNA [19].

We set out to experimentally study the structure of ING3, the least studied of the ING proteins, with the same strategy as previously applied on ING4 and ING5, by NMR, crystallography, and other biophysical techniques. The solubility of ING3 was too low to prepare samples for structural analysis, but its N-terminal domain forms homodimers with antiparallel coiled-coil structure, and the C-terminal domain folds into a PHD that binds histone H3 N-terminal tail, preferentially when trimethylated at K4. These results indicate that ING3 is a bivalent molecule with two PHD that independently recognize H3K4me3.

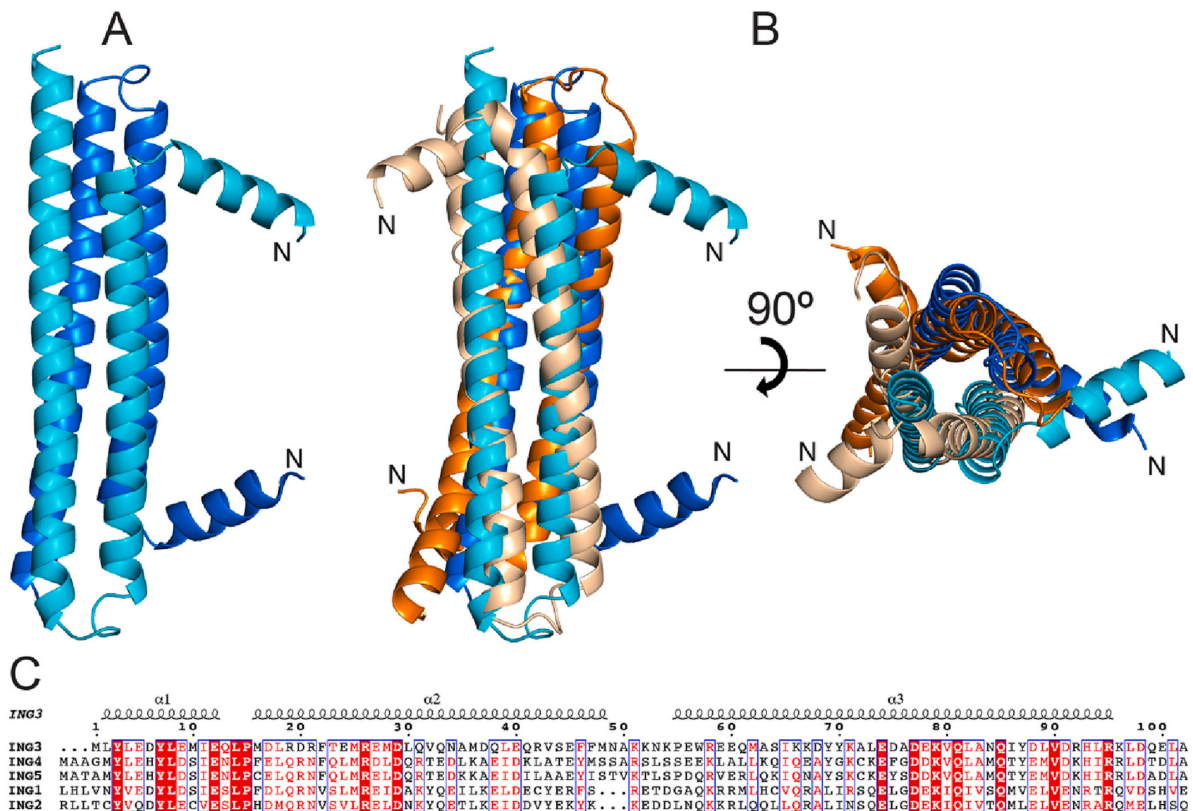
## 2. Results

### 2.1. The N-terminal domain of ING3 is a dimer with coiled-coil structure

Full-length human ING3 protein was produced in bacterial cells as insoluble material, and attempts to recover it in soluble pure form were unsuccessful. The AlphaFold model structure at the Uniprot entry

Q9NXR8 shows a helix-loop-helix forming a coiled-coil at the N-terminal region, a long disordered central region, and a PHD at the C-terminal region. This structural organization is similar to the experimentally studied structures of the homologous ING4 and ING5 proteins [16,17]. However, both ING4 and ING5 form homodimers through their N-terminal domains. The high sequence conservation at this region in ING3-5 (see Fig. 2C below) suggests that ING3 also forms antiparallel coiled-coil homodimers. To verify this hypothesis, we produced and analyzed the structure of the N-terminal domain of ING3 (Nt). SEC-MALS shows a molar mass of 24 kDa, which is double the calculated for a monomer (12.6 kDa), indicating stable homodimer formation (Fig. 1A). The protein elutes at a smaller volume than expected for a globular dimer, suggesting that it has an elongated shape and/or long disordered regions [20]. The circular dichroism spectrum shows a high content of  $\alpha$ -helical structure (Fig. 1B), with an ellipticity ratio  $[\Theta]_{222\text{nm}}/[\Theta]_{208\text{nm}} > 1$ , typical of coiled-coils [21]. The thermal denaturation is cooperative and reversible, with a mid-point temperature of 42 °C (Fig. 1C). Thermal stability is reduced relative to ING4 (51 °C) or ING5 (47 °C) Nt domains under similar conditions. The NMR spectrum shows less backbone amide signals than expected from the amino acid sequence, with low dispersion, and with different line widths (Fig. 1D). These features are consistent with the formation of a symmetric dimer (no duplication of signals) with helical structure (low dispersion), and elongated shape (signal broadening), as observed on the Nt domain of ING4 [16].





**Fig. 2.** Crystal structure of the Nt domain of ING3. (A) Ribbon diagram of the structure with the two protomers colored marine and cyan. (B) Two views of the aligned crystal structures of the Nt domains of ING3 (marine and cyan ribbons) and ING4 (PDB 4AFL, orange and wheat ribbons). The N-terminal ends are indicated in all chains. (C) Alignment of the Nt sequences of human ING proteins with colors and boxes indicating conserved residues and with the secondary structure of ING3 on top, created with ESPRIPT [23]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

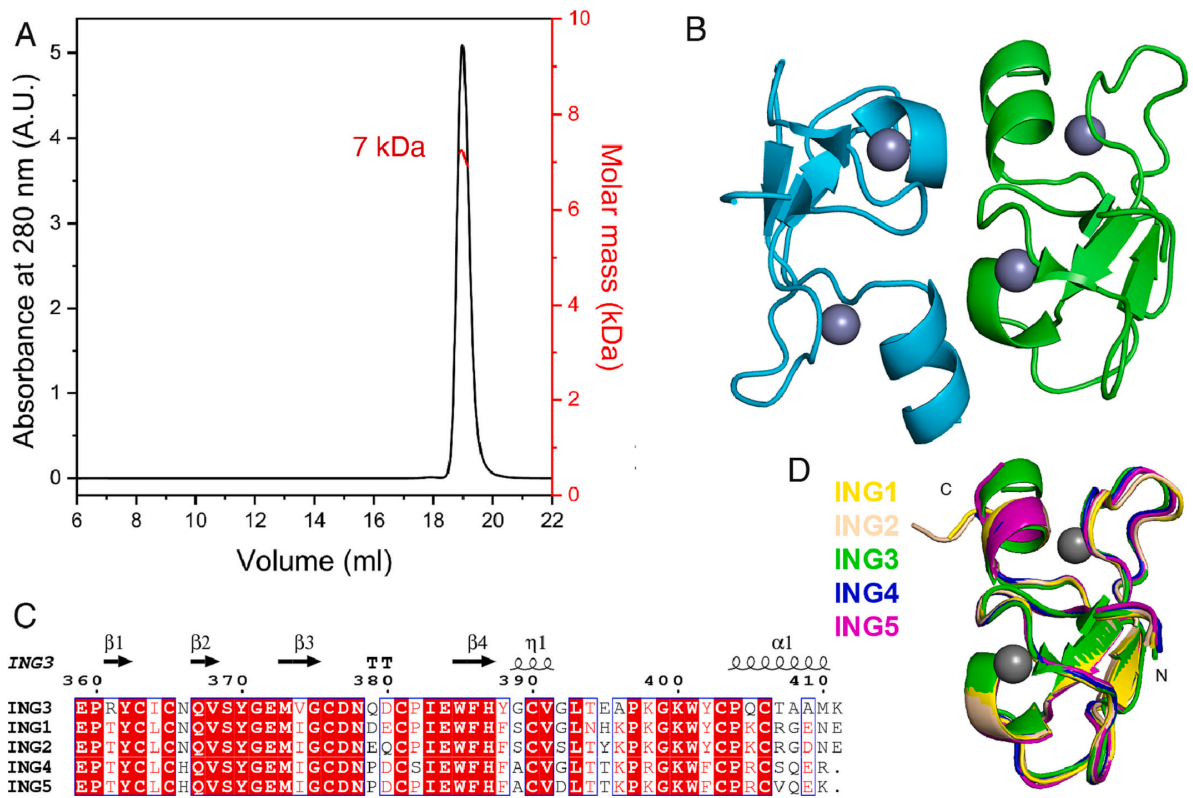
The crystal structure of ING3 Nt (Fig. 2A) shows two chains in the asymmetric unit assembled in an antiparallel coiled-coil. Residues –1 to 97 are observed in chain A and residues –1 to 100 in chain B. Each protomer is folded as a long helix-loop-helix motif, with a short N-terminal helix that is oriented away from the coiled-coil. The program Socket2 [22] identifies the typical knob-into-holes pattern of coiled-coils along 4.1 heptad repeats in each chain. The structure is very similar to that of ING4 (Fig. 2B), the major difference being the orientation of the short N-terminal helix. Excluding this helix, the average pairwise  $\alpha$  RMSD between ING3 and ING4 chains is 1.4 Å. As was found in the crystal structures of ING4 and ING5 Nt domains, the short helix does not form part of the coiled-coil, and makes contacts with other chains in the crystal lattice (for instance, the amide proton of non-native A-1 in chain B is hydrogen-bonded with the carboxyl group of E74 of a nearby symmetry mate). This region very likely is highly dynamic in the three ING homologs. An AlphaFold model of the ING3 Nt dimer (not shown) shows a single long helix at the N-terminal half of the two chains.

## 2.2. Structure of the C-terminal domain of ING3 and histone H3 binding

During the purification of the PHD of ING3 we observed an anomalous migration in reducing SDS-PAGE. While the ubiquitin-PHD fusion migrated predominantly according to its calculated molecular weight (19 kDa), the isolated PHD (6.5 kDa) migrated as a broad band with an apparent molecular weight of approximately 40 kDa. SEC-MALS analysis, however, confirmed that the PHD is monomeric (Fig. 3A), a result

consistent with NMR spectra showing sharp and dispersed signals, typical of a small folded protein (see Fig. 5 below). Crystals diffracted up to 1.2 Å resolution (Table 1) and contained two PHD molecules in the asymmetric unit (Fig. 3B), but the interface is small, with no intermolecular hydrogen bonds or salt bridges. PISA analysis [24] indicates that the interface is not significant (Complexation Significance Score = 0.159). Each molecule shows the typical PHD fold stabilized by two  $\text{Zn}^{2+}$  cations coordinated by the side chains of four cysteines or three cysteines and one histidine. Three  $\text{Ca}^{2+}$  cations contact carboxyl and carbonyl groups of the PHD chains and water molecules, but they do not form part of the PHD fold. The four non-native N-terminal residues and the first native residue (N358) are not observed, and G371 is not visible in chain B. The structure of the two molecules is the same (RMSD is 1.00 Å for  $\text{C}\alpha$  atoms and 1.62 Å for all heavy atoms). The major deviations occur at the N-terminal residues 359–360. The structure is highly similar to the PHD of ING4 and ING5 (RMSD is 0.49 and 0.75 Å, respectively), and to ING1 and ING2 (RMSD is 1.18 and 1.95 Å, respectively), the four of them solved bound to H3K4me3 peptides (Fig. 3C). A solution structure of an ING3 PHD construct is deposited in the Protein Data Bank with entry 1X4I, but is not described in any publication. This construct includes residues 362–418 of ING3 plus seven and six non-native residues at the N and C-terminus, respectively. The overall structure in the well-defined regions of the NMR-ensemble is similar to the crystal structure, with the C-terminal  $\alpha$ -helix elongated by 3 residues.

To examine the binding of the PHD of ING3 to H3K4me3 we tried to crystallize the complex, unsuccessfully. Calorimetry measurements



**Fig. 3.** Crystal structure of the C-terminal domain of ING3. (A) SEC-MALS analysis of ING3 PHD. (B) The two PHD molecules in the asymmetric unit of the crystal (green and cyan ribbons) with the zinc cations (grey spheres). (C) Alignment of the amino acid sequences of the PHDs of the five human ING proteins extracted from their crystal structures chains A of the PDB entries 2QIC, 2G6Q, 7ZMX, 2VNF, 3C6W) with the ING3 secondary structure elements on top ( $\beta$ -strand, turn,  $3_{10}$  helix and  $\alpha$ -helix, in order of appearance from left to right). Residue numbering corresponds to the PHD of ING3. (D) Superposition of the crystal structures of the five PHDs. The N- and C-termini are indicated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

confirmed the binding to H3K4me3 ( $K_D = 2.5 \pm 0.3 \mu\text{M}$  at  $25^\circ\text{C}$ ), with smaller affinity to H3K4me2 and H3K4me1, and undetectable (under the conditions used) for H3K4me0 (Fig. 3). These affinities are very similar to those measured by tryptophan fluorescence at  $25^\circ\text{C}$  and by calorimetry at  $5^\circ\text{C}$  [13,25]. The binding is enthalpy driven (Table 2), as found before for the PHD of ING4 [14].

To characterize the structural basis of the binding by NMR, we assigned the spectrum of the PHD and then titrate it with the H3K4me3 peptide (Fig. 5A). Several residues were perturbed, some of them experiencing large shifts, and a few of them also showing strong signal broadening at sub-stoichiometric protein:peptide molar ratios. These observations indicate binding occurring in the fast-intermediate exchange regime on the chemical shift time scale. Docking of the histone peptide to the PHD, guided by the perturbations seen in the NMR spectrum, shows the peptide bound antiparallel to one of the strands of the short  $\beta$ -sheet of the PHD (Fig. 5B). The trimethylated amino group of K4 in the peptide forms cation- $\pi$  interactions with the side chains of Y362 and W385, and the carbonyl of G399 interacts with the amino group of A1 in the peptide (see below). These and other features are also observed in the complexes of the other ING homologs, indicating that the basis for the molecular recognition is very similar in all of them. The NMR titration of the PHD with the H3K4me0 peptide shows perturbations occurring in the fast exchange regime, with little or no broadening (Fig. 5C). Mapping on the PHD structure the most perturbed residues shows that the binding site is the same as for H3K4me3. However, the affinity is 54-fold lower ( $K_D = 134 \pm 8 \mu\text{M}$  at  $25^\circ\text{C}$ , derived from fitting

of the CSPs to a one site binding model; Fig. 5D).

The flexibility of the backbone of the PHD free and bound to H3K4me3 peptide was probed by measuring the  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOEs, which are sensitive to the local main chain dynamics in the ps-ns time scale (Fig. 6). Values smaller than 0.65 are considered flexible, and the smaller the value the larger the flexibility. With this criterion, the flexible PHD residues are the first two N-terminal (native) residues, the C-terminal residue, and G399, in one of the loops. Upon H3K4me3 binding the values are very similar for most of the residues, but notably larger for G399. This could be explained by the interaction between the carbonyl of G399 and the amino group of A1 of the histone peptide (Fig. 6).

### 3. Discussion

In contrast with ING4 and ING5, we were unable to prepare soluble samples of full-length ING3 for structural analysis. The most likely reason is that the central disordered region of ING3 is three times longer (256 residues) and contains the same number of acidic and basic residues. The central regions of ING4 and ING5 have excess of basic residues, favoring solubility. Still, ING3 has 24 positively charged residues in this region, and possibly binds dsDNA, as ING4 and ING5 do. The structures of the isolated folded regions indicate that ING3 forms antiparallel dimers by its Nt domain and has two PHD that independently bind H3K4me3. Sequence homology is lower for ING1 and ING2, but it is likely that they also form homodimers [18], and that the five ING

**Table 1**  
Diffraction data collection and crystal structure refinement statistics of ING3 domains.

	PHD	Nt
Data collection		
Space group	P3 <sub>1</sub> 21	I432
Cell dimensions		
a, b, c (Å), α, β, γ (°)	40.69, 40.69, 100.59	149.39, 149.39, 149.39
	90, 90, 120	90, 90, 90
Wavelength (Å)	0.979182	0.97926
Resolution (Å) <sup>a</sup>	35.24–1.20 (1.22–1.20)	74.69–2.91 (3.00–2.91)
Reflexions (unique)	423,848 (47743)	464,180 (6277)
CC (1/2) (%)	99.9 (75.9)	100 (56.5)
R <sub>merge</sub> (%)	8.5 (61.3)	8.0 (660.9)
R <sub>meas</sub> (%)	9.4 (72.3)	8.1 (665.1)
R <sub>pim</sub> (%), separate Friedels	3.2 (41.4)	1.0 (73.8)
Completeness (%)	92.4 (45.6)	95.7 (55.9)
<I/σ(I)>	18.6 (0.2)	41.4 (1.0)
Multiplicity	14.8 (5.5)	73.9 (80.3)
Wilson B-factor	15.70	126.07
Refinement		
Resolution (Å)	35.24–1.20	74.69–2.91
No. reflections	44,362	6276
R <sub>work</sub> / R <sub>free</sub>	17.19 (43.04) / 20.47 (44.81)	26.96 (46.19) / 31.54 (54.16)
No. atoms		
Protein	829	1702
Water	98	–
Zn <sup>2+</sup> ions	4	–
Ca <sup>2+</sup> ions	3	–
B factors (Å <sup>2</sup> )		
Protein atoms	29.52	126.73
Water molecules	33.37	–
Zn <sup>2+</sup> ions	22.33	–
Ca <sup>2+</sup> ions	27.98	–
R.m.s. deviations		
Bond lengths (Å)	0.009	0.009
Bond angles (°)	1.000	1.310
Ramachandran statistics		
Preferred	93.94 %	90.26 %
Allowed	6.06 %	8.72 %
Disallowed	0.00 %	1.03 %
PDB ID	7ZMX	8COK

<sup>a</sup> Statistics for the highest-resolution shell are shown in parentheses.

**Table 2**  
Thermodynamic parameters of histone H3 peptide binding to ING3 PHD at 25 °C.

	K <sub>d</sub> <sup>a</sup> (μM)	ΔG <sup>b</sup> (kcal/ mol)	ΔH <sup>c</sup> (kcal/ mol)	–TΔS <sup>d</sup> (kcal/ mol)	N <sup>e</sup>
H3K4me0 <sup>b</sup>	134 ± 7	–	–	–	1
H3K4me1	13.8 ± 0.2	–6.8	–7.4 ± 1.3	0.6	1.08 ± 0.08
H3K4me2	5.9 ± 0.9	–7.1	–7.9 ± 0.4	0.7	1.09 ± 0.03
H3K4me3	2.5 ± 0.2	–7.7	–10.6 ± 0.3	2.9	0.94 ± 0.01

The errors are fitting errors. Nd: not determined.

<sup>a</sup> Dissociation constant.

<sup>b</sup> Measured by NMR (average of nine residues, and average of the standard errors).

<sup>c</sup> Standard free energy change of the association reaction.

<sup>d</sup> Standard enthalpy change of the association reaction.

<sup>e</sup> Standard entropy change of the association reaction multiplied by the absolute temperature.

<sup>f</sup> Stoichiometry, which was an adjustable parameter in the fitting of the calorimetry data, or fixed to 1 in the case of NMR data.

proteins are bivalent readers of the histone H3K4me3 mark.

Although ING3 is not frequently mutated in tumors, our crystal structures predict that some missense mutations described in the Catalog Of Somatic Mutations In Cancer (COSMIC) will reduce ING structural stability. Mutations F46L and L94F may destabilize the Nt domain because both residues are buried inside the hydrophobic core, and may destabilize homodimer formation because the two residues are in contact at opposite ends of the antiparallel coiled-coil. Mutation G375R will strongly destabilize the PHD because of the introduction of a large and charged side chain in a buried position. And C381F and C390F mutations will strongly destabilize the PHD as these cysteines are involved in the coordination of the two Zn<sup>2+</sup> cations. No other missense mutations described in tumors are expected to have a strong impact on H3K4me3 recognition by ING3. However, our structure explains previous site-directed mutagenesis studies showing that Y362A reduces affinity for H3K4me3, and W385A shows no binding [25,26]. The two aromatic residues form cation-π interactions with the trimethylammonium of K4me3 (Fig. 6).

The affinity of the five isolated PHD for histone H3K4me3 is in the low-micromolar range [13], but macromolecular crowding increases the affinity of ING4 PHD by one order of magnitude [27], bivalency increments the affinity 3-fold for ING4 and ING5 [16,17], and dsDNA binding affinity is in the low-micromolar range for ING4 [19]. Therefore, the affinity of ING proteins for the nucleosomes in the nucleus is probably in the nanomolar range.

Inside the cell ING proteins associate with other proteins into HAT or HDAC complexes [7]. ING3 is part of the Tip60 complex, containing the Tip60 HAT (of the MYST family) and at least 15 other proteins [28]. It is conserved from yeast to humans, and in yeast it is called the NuA4 complex, with 13 subunits, including the HAT Esa1 and the Yng2 protein, the closest ING3 homolog in yeast. The structure of the Tip60 complex is not known, but the structure of NuA4 bound to the nucleosome has been recently determined by cryoEM [29]. It is organized in a HAT module and a transcription activator-binding module. The HAT module is organized into the catalytic core and the helical bundle. The helical bundle shows a four-helix antiparallel coiled coil formed by the N-terminal region of Yng2 folded into a helix-loop-helix structure, a long helix of protein Eaf6, and a long helix of protein Epl1. The C-terminal PHD and the central region of Yng2 are not observed, likely due to the central region being disordered and the PHD not binding the histones H3 tails with non-methylated K4. The structure of the Yng2 protein is the same as observed in a NuA4 sub-complex cryoEM structure [30] and in the crystal structure of a four-subunit subcomplex [31], where only the N-terminal domain of Yng2 was used for crystallization. Whether the Tip60 complex contains only one ING3 molecule, as happens with its homolog Yng2 in the NuA4 complex, or two ING3 molecules, as our results suggest, remains to be determined. The bivalency of dimeric ING3 would enhance the targeting of the Tip60 HAT complex to chromatin sites enriched in the histone H3K4me3 mark.

## 4. Materials and methods

### 4.1. Gene construct design for protein production

The gene construct for full-length human ING3 protein production (UniProt entry Q9NXR8, isoform 1) was designed with an N-terminal fusion to ubiquitin to favor gene expression and protein solubility in bacterial cultures [32]. Because the human gene contains codons that are rare in *E. coli*, a synthetic gene with optimized codons was purchased from Genscript. The ubiquitin moiety contains a His<sub>10</sub>-insertion in the loop between the first β-hairpin and the helix to facilitate protein purification by immobilized metal affinity chromatography (IMAC). Ubiquitin was followed by the TEV-protease cleavage sequence to remove the fusion protein. Between the TEV site and the initial methionine of ING3 an alanine was inserted to facilitate TEV cleavage. This construct was cloned in plasmid pET29a(+) under the control of T7 promoter. The

gene constructs for the N-terminal and the C-terminal domains of ING3 were designed and cloned in the same way. The high sequence homology at the N-t and C-terminal regions of ING3–5 was used to define the boundaries of the folded domains, based on those experimentally identified in ING4 and ING5 [16,17]. The Nt domain comprises residues M1-A102, and the PHD residues N358-K411. The last seven residues of ING3 are non-conserved, charged or polar, predicted to be disordered, and were excluded to facilitate PHD crystallization. The Nt protein has, after purification, the non-native GA sequence preceding its N-terminus. In the case of PHD, the non-native sequence at its N-terminus is GAMG. The two extra residues were included to have a polypeptide chain similar to the ING4 PHD that was successfully co-crystallized with H3K4me3 peptide.

#### 4.2. Protein production and purification

*E. coli* BL21(DE3) cells harboring the corresponding plasmids were grown in peptone-yeast extract-NaCl medium or auto-induction ZYP-5052 medium [33] for natural isotopic abundance, and in P-50501 auto-induction media [34] for uniform  $^{13}\text{C}$ ,  $^{15}\text{N}$  isotopic enrichment. For full-length ING3 and the PHD, the media were supplemented with 50  $\mu\text{M}$   $\text{ZnCl}_2$ , because two  $\text{Zn}^{2+}$  cations are necessary for PHD folding, and all media contained 30 mg/l kanamycin. Cells in peptone-based medium were grown at 37 °C until  $\text{OD}_{600} = 0.6$ , when expression was induced with 1 mM IPTG for 16–20 h at 20 °C. Cells in auto-induction media started at  $\text{OD}_{600} = 0.05$  (from a saturated culture in peptone-based medium) and were grown for 2 h at 37 °C before lowering the temperature to 20 °C for further growth and protein production during 16–20 h. Cells were harvested by centrifugation and resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 1 mM DTT with *Complete EDTA-free* protease inhibitors, Merck). After sonication (with a small amount of lysozyme and DNase) and ultracentrifugation, proteins were found predominantly in the insoluble (ING3), soluble (Nt domain) or in both fractions (PHD). Insoluble proteins were recovered from the pellets by solubilization in 7–8 M urea and ultracentrifugation, and then refolded by a drop-by-drop 50-fold dilution into cold 20 mM, Tris-HCl pH 8.0, 300 mM NaCl, 1 mM DTT plus 50  $\mu\text{M}$   $\text{ZnCl}_2$  with mild agitation. The proteins were purified by IMAC on a 5 ml His-Trap column (Cytiva) loaded with  $\text{Ni}^{2+}$  and equilibrated in 20 mM Tris-HCl pH 8.0, 300 mM NaCl, and 1 mM DTT. After washing with buffer with 25 mM imidazole, the protein was eluted with 500 mM imidazole and dialyzed for 15–20 h at 4 °C against the buffer without imidazole at the same time that was cleaved by addition of His-tagged TEV protease (1 mg of protease per 30 mg of total protein). The cleavage was complete, partial, or difficult to assess for the Nt domain, the PHD, and ING3, respectively. The samples were loaded on an equilibrated His-Trap column to collect the cleaved protein in the flow-through. For ING3 this procedure yielded no detectable protein in Coomassie-blue stained SDS-PAGE. For the N-terminal domain, further purification was done by anion exchange and size exclusion chromatography (SEC). After diluting the protein solution to reduce NaCl concentration to 40 mM NaCl, it was loaded on a 5 ml Hi-Trap Q HP (Cytiva) equilibrated in 20 mM Tris-HCl pH 8.0, 40 mM NaCl, and 1 mM DTT, and eluted with a linear gradient until 1 M NaCl. The protein containing fractions were concentrated and loaded on a Superdex 75 26/60 in 20 mM Tris-HCl pH 8.0, 40 mM NaCl, and 1 mM DTT. For the PHD, the flowthrough of the second IMAC was polished by SEC on a Superdex 75 16/60. The purity, concentration, and identity of the purified proteins were confirmed by Coomassie-stained SDS-PAGE, ultraviolet absorbance spectroscopy, and MALDI-TOF mass spectrometry. ING3 and N-t domain proteins were analyzed in 12 % acrylamide/bisacrylamide (29:1, by weight) gels run with Tris/Gly buffer. The small PHD protein was analyzed in 10 % gels run with Tris/Tricine buffer. PHD purified from the soluble or the insoluble fractions of lysed cells yielded the same folded protein, with identical NMR and mass spectra. Pure proteins were concentrated by ultrafiltration and the concentration was measured by absorbance at 280 nm using the extinction coefficient

calculated from the amino acid sequence on the ExPASy ProtParam server [35]. Aliquots of the pure proteins were frozen in liquid nitrogen and stored at  $-80$  °C until used. When necessary, the aliquots were thawed on ice and buffer exchanged by three cycles of dilution-concentration using Amicon (Merck) centrifugal filters.

#### 4.3. Peptides

Synthetic lyophilized peptides with free termini were purchased from PolyPeptide or from Apeptide. Four peptides correspond to histone H3 residues 1–15 plus a Tyr residue at the C-terminus (ARTKQTARKSTGGKAY) with the four possible methylation states at K4. Solution stocks were prepared by dissolving powder in PBS (10 mM phosphate, 140 mM chloride, 153 mM sodium, and 4.5 mM potassium, pH 7.4) and the pH was adjusted to 7.0 with concentrated NaOH and HCl. Peptide concentration was measured by absorbance at 280 nm using the extinction coefficient of the tyrosine. A peptide corresponding to histone H3 residues 1–10 (ARTKQTARKS) trimethylated at K4 was used for (unsuccessful) crystallization trials.

#### 4.4. Circular dichroism (CD)

The spectrum of ING3 Nt was measured on a Jasco J-715 spectropolarimeter at 25 °C on a 35  $\mu\text{M}$  sample in PBS at pH 7.4 in a 0.1 mm path length quartz cuvette, and is the average of 10 scans and base-line corrected. Thermal denaturation was measured in a 2 mm cuvette closed with a Teflon cap on a 15  $\mu\text{M}$  sample. Denaturation was monitored at 222 nm by increasing temperature at 1 °C/min from 10 °C to 90 °C. The midpoint of the denaturation curve was determined from its first derivative. Reversibility was examined by comparing spectra at 25 °C on the same sample before and after denaturation.

#### 4.5. Size exclusion chromatography-multiangle light scattering (SEC-MALS)

The experiments on the PHD of ING3 were performed at 25 °C using a Superdex 200 10/300 GL column (Cytiva) connected to a DAWN-HELEOS light scattering detector and an Optilab rEX differential refractive index detector (Wyatt Technology). The column was equilibrated with PBS at pH 7.0, 0.5 mM TCEP (0.1  $\mu\text{M}$  filtered). A sample of 100  $\mu\text{l}$  protein at 1 g/l was injected and run at 0.5 ml/min. For the Nt domain, 200  $\mu\text{l}$  at 1.5 g/l were injected in the same column running in PBS pH 7.4, and the light dispersion was measured at 23 °C with a DAWN-EOS detector. Data acquisition and analysis was done with ASTRA software (Wyatt). The SEC-MALS systems were calibrated with a sample of Bovine Serum Albumin (BSA) at 1 g/l in the same buffer. Based on numerous measurements on BSA under the same or similar conditions, we estimate that the experimental error in molar mass is around 5 %.

#### 4.6. NMR spectroscopy

The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of the N-t domain was recorded 25 °C on a Bruker AVANCE 600 spectrometer with a TXI cryoprobe and z-gradient coil. A 150  $\mu\text{M}$  sample in 20 mM sodium phosphate pH 6.5, 5 %  $^2\text{H}_2\text{O}$ , 50  $\mu\text{M}$  DSS (2,2-dimethyl-2-silapentane-5-sulfonate sodium salt) was used. Triple resonance NMR spectra of the PHD were recorded at 25 °C on a Bruker AVANCE III 600 spectrometer with a z-gradient coil probe. A  $^{13}\text{C}$ ,  $^{15}\text{N}$  labeled ING3 PHD protein sample was prepared at 920  $\mu\text{M}$  in 20 mM sodium phosphate pH 6.5, 50 mM NaCl, 2 mM DTT, 5 %  $^2\text{H}_2\text{O}$ , 0.01 %  $\text{NaN}_3$ , 50  $\mu\text{M}$  DSS. Backbone  $\text{H}^{\text{N}}$ ,  $\text{H}_\alpha$ , N, CO, C $\alpha$ , and C $\beta$  assignments were obtained from 3D HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HN(COCA)HA and HN(CA)HA experiments implemented as BEST-TROSY [36] and acquired with non-uniform sampling [37]. Chemical shifts were measured relative to internal DSS for  $^1\text{H}$  and calculated for  $^{13}\text{C}$  and  $^{15}\text{N}$  [38]. NMR data were

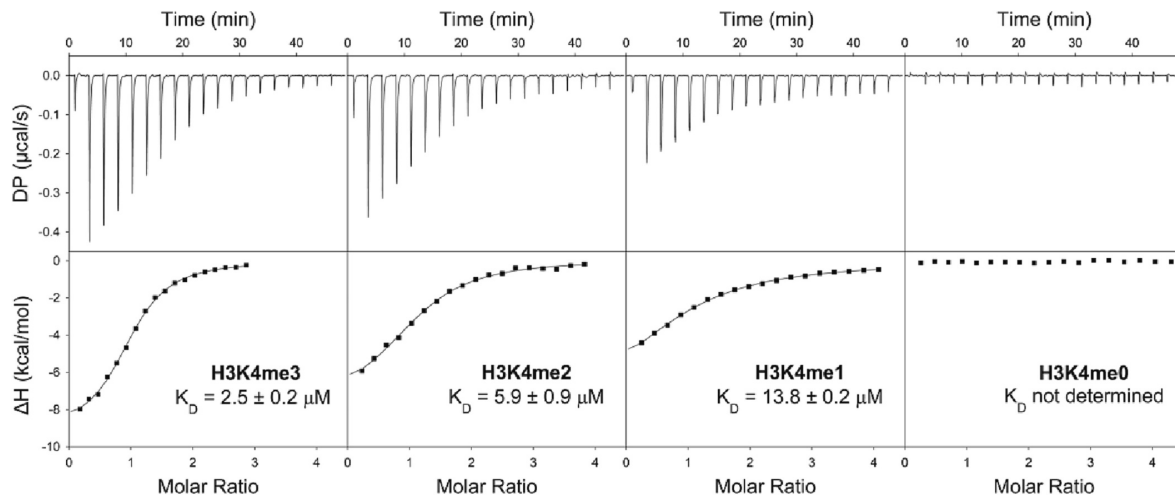


Fig. 4. ING3 PHD binding to histone H3 peptides studied by ITC. The upper panels illustrate the heat change caused by PHD binding to the different peptides. The lower panels show the integration of each peak plotted against the molar ratio peptide:PHD after normalization and correction for the heat of dilution. The symbols correspond to the experimental data per injection, and the continuous line corresponds to the best fit to a model of one set of identical binding sites.

processed with TopSpin (Bruker) and analyzed with NMRFARM-Sparky [39]. The assignment was initially done automatically using MARS [40] and manually completed. The NMR assignment has been deposited in the BioMagResBank entry 51196. There is a deposited assignment of a different construct of the ING3 PHD with entry 11353, which has not been described in a publication. Some of the amide backbone assignments differ from ours. Our assignment is consistent with the labels shown in a reported  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of a similar construct of the ING3 PHD [25], but this assignment is not deposited in the BioMagResBank. On the same sample used for the assignment, the heteronuclear  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE was measured, in an interleaved mode, with a saturation time of 3 s and a relaxation delay of 6 s. The titration of the PHD with the H3K4me3 peptide was performed on Bruker AVANCE III 800 spectrometer with a TCI cryoprobe and z-gradient coil. Small volumes of the peptide stock (4 mM of H3K4me3 in 20 mM sodium phosphate pH 6.5, 50 mM NaCl, 1 mM DTT) were pipetted inside a shigemi tube (without plunger) containing 400  $\mu\text{l}$  of 16.5  $\mu\text{M}$  PHD in the same buffer, and mixed by capping and inverting the tube several times.  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were measured after each addition. The assignment of the PHD signals bound to H3K4me3 was confirmed with 3D-HNCO, HNCA, and HN(CA)CO spectra recorded on a 220  $\mu\text{M}$  PHD sample with 880  $\mu\text{M}$  H3K4me3 peptide. This assignment has been deposited in the BioMagResBank entry 51,225. On this same sample, the heteronuclear  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE was measured in the same way as for the isolated PHD. The titration of the PHD with the H3K4me0 peptide was performed in the same way (with a 3.3 mM peptide stock) on a Bruker AVANCE 600 spectrometer with a TXI cryoprobe and z-gradient coil.

The chemical shift perturbations (CSP) of the PHD signals caused by peptide binding were computed as the weighted average distance between the backbone amide  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts in the bound and free states ( $\text{CSP} = ((\Delta\delta_{\text{H}})^2 + (\Delta\delta_{\text{N}}/5)^2)^{1/2}$ ), as described [41]. The estimated CSP error is  $\pm 0.005$  ppm. The dissociation constant for the PHD-H3K4me0 complex was derived from those residues with a CSP larger than the average plus one standard deviation at the last point of the titration. The CSP values along the titration were fitted to a 1:1 binding model for each residue. The reported  $K_{\text{D}}$  value is the mean, and the error is the mean of the standard errors.

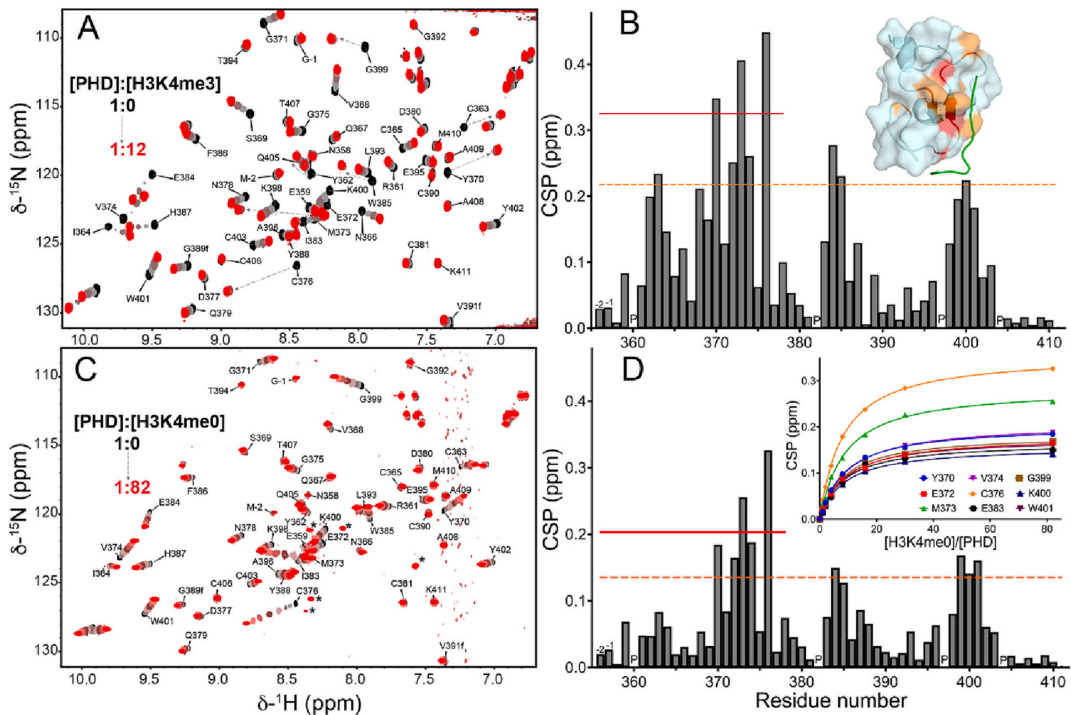
#### 4.7. Crystallization, data collection and structure determination

The PHD and the Nt domain of ING3 protein were dialyzed against

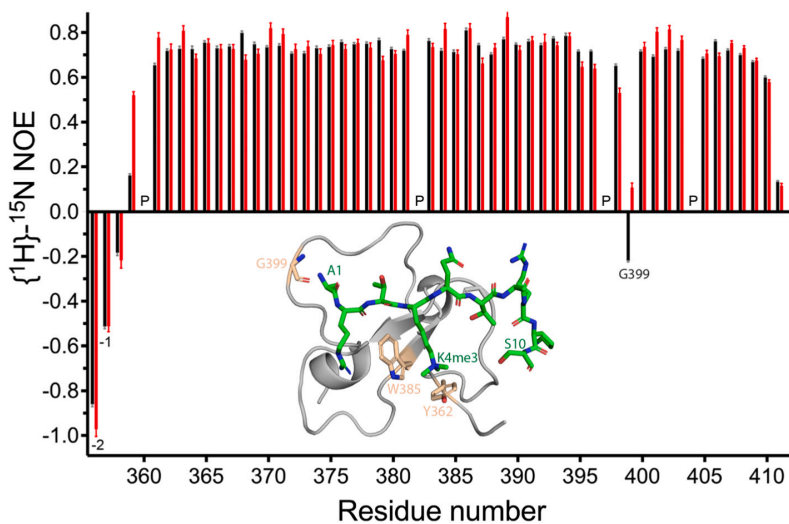
20 mM Tris-HCl pH 6.5, 150 mM NaCl, and 1 mM DTT and concentrated to approximately 3 mM (20 g/l, PHD) and 1.3 mM (16 g/l, Nt domain). Initial crystallization trials were performed at 22  $^{\circ}\text{C}$  using the sitting-drop vapor-diffusion method with commercial screening solutions, including JBScreen Classic and Wizard Classics I-IV (Jena Bioscience, Jena, Germany), in 96-well sitting-drop plates (Swissci MRC; Molecular Dimensions, Suffolk, England). Drops were set up by mixing equal volumes (0.2  $\mu\text{l}$ ) of protein-containing solution and reservoir solution using a Cartesian Honeybee System (Genomic Solutions, Irvine, USA) nanodispenser robot and equilibrated against 50  $\mu\text{l}$  reservoir solution. Crystals of PHD grew in three days in 0.1 M Tris-HCl at pH 8.4, 30 % PEG 4000, and 0.2 M  $\text{CaCl}_2$ . The same crystallization trials were conducted in the presence of H3K4me3 at 1:1.5 or 1:2.8 PHD:peptide molar ratios, unsuccessfully. For data collection, crystals were harvested in the precipitant solution containing 30 % PEG 400 and flash-cooled in liquid nitrogen. Small crystals of the Nt domain grew in 0.1 M sodium acetate buffer at pH 4.6, 30 % 2-Methyl-2,4-pentanediol (v/v) and 20 mM  $\text{CaCl}_2$ . These crystals were used as seeds for new crystallization assays with 5 g/l protein. Larger crystals grew in one week in 0.1 M phosphate-citrate buffer at pH 4.2, 10 % 2-propanol, and 0.2 M  $\text{Li}_2\text{SO}_4$ . For data collection, crystals were harvested in the precipitant solution containing 30 % glycerol and flash-cooled in liquid nitrogen. X-ray diffraction data were collected at the ALBA Synchrotron (Cerdanyola del Vallès) BL13 XALOC beamline. Data were indexed, integrated, scaled, and merged with XDS [42]. Data collection details and refinement statistics are in Table 1. The structure of PHD was solved by molecular replacement using the crystal structure of the PHD of ING4 (PDB 2VNF) with Phaser [43]. The structure of the Nt domain was solved by molecular replacement using the crystal structure of the Nt of ING4 (PDB 4AFL) with MrBUMP [44]. For both domains, the initial model was first refined with Phenix [45] and alternating manual building with Coot [46]. The final models were obtained by repetitive cycles of refinement; solvent molecules were added automatically and inspected visually for chemically plausible positions. The stereochemical quality of the model was assessed with MolProbity [47]. The coordinates of both domains have been deposited with the PDB (accession code 7ZMX for the PHD and 8COK for the Nt domain).

#### 4.8. Isothermal titration calorimetry

Measurements were performed on a MicroCal PEAQ-ITC calorimeter



**Fig. 5.** ING3 PHD binding to histone H3 peptides studied by NMR. (A) Overlay of  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of PHD at 25 °C in the absence (black) or presence (from grey to red) of increasing amounts of H3K4me3 peptide. At the last point of the titration 99 % of PHD is calculated to be bound to peptide. Backbone amide signals are labeled, and those with an “f” are folded in the  $^{15}\text{N}$  dimension. All spectra are plotted with the same contour levels. (B) Bar graph of the CSP for PHD residues measured at the last point of the titration. The non-native residues M-2 and G-1 are indicated. The position of proline residues is indicated by the letter “P”. The average CSP plus one standard deviation is represented by the dashed orange line, and average plus two standard deviations by the red continuous line. The inset is a transparent surface representation of the PHD backbone structure (seen as a ribbon) with the corresponding residues colored orange or red. The backbone of the docked peptide is shown in green. (C) Same as in A for the titration with H3K4me0 peptide. Signals labeled with an asterisk appear at late titration points, likely natural  $^{15}\text{N}$  abundance signals from the excess free peptide (at 1.4 mM at the last titration point). At the last point of the titration 88 % of PHD is calculated to be bound to peptide. (D) Same as in B for H3K4me0 peptide. The inset shows the fitting to a 1:1 binding model of the CSP values along the titration for the indicated residues. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).



**Fig. 6.** Backbone flexibility of ING3 PHD. Heteronuclear  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOEs for the backbone amides of the PHD of ING3 free (black) or bound to histone H3K4me3 peptide (red) with error bars. The asterisks indicate the two non-native residues M-2 and G-1 preceding N358. The position of proline residues is indicated by the letter “P”. The inset shows the structure of the histone H3K4me3 peptide bound to ING4 PHD (sticks with green for carbon, blue for nitrogen, and red for oxygen atoms) docked on the crystal structure of ING3 PHD (grey ribbon, with main chain and side chains of G399, W385 and Y362 in sticks with carbons colored wheat). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

(Malvern) at 25 °C in PBS pH 7.0, 1 mM DTT. The protein and the peptides were separately dialyzed against the same buffer. The sample cell contained 20 μM PHD and the syringe contained peptides at 300 μM (H3K4me3) or 400 μM (H3K4me2, 1 and 0). The experiment consisted of a series of 13 injections of 3 μl or 19 injections of 2 μl, with a 150 s delay and stirring at 750 rpm. The electrical power required to maintain the cell at constant temperature after each injection was recorded as a function of time, generating the corresponding thermograms. Data were fitted to a 1:1 binding model using the MicroCal PEAQ-ITC software (Malvern). A few trials with different protein and peptide concentrations as well as injection patterns were done to find well-performing conditions, but the measurements shown in Fig. 4 were performed only once.

#### 4.9. Molecular modelling

The Hdock server (<http://hdock.phys.hust.edu.cn/>) was used to dock the H3K4me3 peptide to the ING3 PHD [48]. Chain B of the PHD crystal structure was defined as the receptor, and the structure of the H3K4me3 peptide bound to the PHD of ING4 (chain D in PDB 2VNF) was defined as the ligand. PHD residues with CSP larger than the average plus two standard deviations (Y370, M373, and C376) were defined as the binding site. The model with the highest quality score (out of 10 models) was chosen as representative of the complex. H3K4me0 was docked defining M373 and C376 as residues at the binding site.

Figures were generated with PyMol (<http://www.pymol.org>) and with the ESPRIPT web server (<http://escript.ibcp.fr/EScript/EScript/>).

#### CRedit authorship contribution statement

**Mariola Ferreras-Gutiérrez:** Investigation, **Belén Chaves-Arquero:** Investigation **Amaia González-Magana:** Investigation **Nekane Merino:** Investigation **Ignacio Amusatategui-Mateu:** Investigation **Sonia Hucas:** Methodology **Francisco J. Medrano:** Investigation **Francisco J. Blanco:** Conceptualization, Writing Original draft preparation, Reviewing and Editing, Project Administration, Funding acquisition.

#### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Francisco J Blanco reports financial support was provided by Spanish Ministry of Science and Innovation.

#### Acknowledgements

We thank Tammo Diercks for help with NMR experiments and Antonio Romero for crystallographic resources. FJB and MF-G were supported by grants PID2020-113225GB-I00 and PRE2018-085788 funded by MCIN/AEI/10.13039/501100011033. AG-M. acknowledges the financial support received from the Spanish Ministry of Universities and the Grants for the requalification of the Spanish university system for 2021-2023, funded by the European Union-Next Generation EU-Margarita Salas Modality.

#### References

















- [1] P.Y. Kan, T.L. Caterino, J.J. Hayes, The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays, *Mol. Cell Biol.* 29 (2) (2009) 538–546.
- [2] A. Groth, W. Rocha, A. Verreault, G. Almouzni, Chromatin challenges during DNA replication and repair, *Cell* 128 (4) (2007) 721–733.
- [3] B. Li, M. Carey, J.L. Workman, The role of chromatin during transcription, *Cell* 128 (4) (2007) 707–719.
- [4] S.D. Taverna, H. Li, A.J. Ruthenburg, C.D. Allis, D.J. Patel, How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers, *Nat. Struct. Mol. Biol.* 14 (11) (2007) 1025–1040.
- [5] G.H. He, C.C. Helbing, M.J. Wagner, C.W. Sensen, K. Riabowol, Phylogenetic analysis of the ING family of PHD finger proteins, *Mol. Biol. Evol.* 22 (1) (2005) 104–116.

- [6] S. Park, G.W. Kim, S.H. Kwon, J.S. Lee, Broad domains of histone H3 lysine 4 trimethylation in transcriptional regulation and disease, *FEBS J.* 287 (14) (2020) 2891–2902.
- [7] Y. Doyon, C. Cayrou, M. Ullah, A.J. Landry, V. Cote, W. Selleck, W.S. Lane, S. Tan, X.J. Yang, J. Cote, ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation, *Mol. Cell* 21 (1) (2006) 51–64.
- [8] M. Nagashima, M. Shiseki, K. Miura, K. Hagiwara, S.P. Linke, R. Pedoux, X. W. Wang, J. Yokota, K. Riabowol, C.C. Harris, DNA damage-inducible gene p33ING2 negatively regulates cell proliferation through acetylation of p53, *Proc. Natl. Acad. Sci. U. S. A.* 98 (17) (2001) 9671–9676.
- [9] M. Shiseki, M. Nagashima, R.M. Pedoux, M. Kitahama-Shiseki, K. Miura, S. Okamura, H. Onogi, Y. Higashimoto, E. Appella, J. Yokota, C.C. Harris, p29ING4 and p28ING5 bind to p53 and p300, and enhance p53 activity, *Cancer Res.* 63 (10) (2003) 2373–2378.
- [10] K. Jaquet, O. Binda, ING proteins: tumour suppressors or oncoproteins, *Cancers (Basel)* 13 (9) (2021).
- [11] X. Shi, T. Hong, K.L. Walter, M. Ewalt, E. Michishita, T. Hung, D. Carney, P. Pena, F. Lan, M.R. Kaadige, N. Lacoste, C. Cayrou, F. Davrazou, A. Saha, B.R. Cairns, D. E. Ayer, T.G. Kutateladze, Y. Shi, J. Cote, K.F. Chua, O. Gozani, ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression, *Nature* 442 (7098) (2006) 96–99.
- [12] P.V. Peña, R.A. Hom, T. Hung, H. Lin, A.J. Kuo, R.P. Wong, O.M. Subach, K. S. Champagne, R. Zhao, V.V. Verkhusha, G. Li, O. Gozani, T.G. Kutateladze, Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor, *J. Mol. Biol.* 380 (2) (2008) 303–312.
- [13] P.V. Peña, F. Davrazou, X. Shi, K.L. Walter, V.V. Verkhusha, O. Gozani, R. Zhao, T. G. Kutateladze, Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2, *Nature* 442 (7098) (2006) 100–103.
- [14] A. Palacios, I.G. Munoz, D. Pantoja-Uceda, M.J. Marcaida, D. Torres, J.M. Martín-García, I. Luque, G. Montoya, F.J. Blanco, Molecular basis of histone H3K4me3 recognition by ING4, *J. Biol. Chem.* 283 (23) (2008) 15956–15964.
- [15] K.S. Champagne, N. Saksouk, P.V. Peña, K. Johnson, M. Ullah, X.J. Yang, J. Cote, T. G. Kutateladze, The crystal structure of the ING5 PHD finger in complex with an H3K4me3 histone peptide, *Proteins* 72 (4) (2008) 1371–1376.
- [16] A. Palacios, A. Moreno, B.L. Oliveira, T. Rivera, J. Prieto, P. García, M. R. Fernandez-Fernandez, P. Bernardo, F.J. Blanco, The dimeric structure and the bivalent recognition of H3K4me3 by the tumor suppressor ING4 suggests a mechanism for enhanced targeting of the HBO1 complex to chromatin, *J. Mol. Biol.* 396 (4) (2010) 1117–1127.
- [17] G. Ormazá, J.A. Rodríguez, A. Ibanez de Opakua, N. Merino, M. Villate, I. Gorrono, M. Rabano, I. Palmero, M. Vilaseca, R. Kypka, M.D.M. Vivanco, A.L. Rojas, F. J. Blanco, The tumor suppressor ING5 is a dimeric, bivalent recognition molecule of the histone H3K4me3 mark, *J. Mol. Biol.* 431 (12) (2019) 2298–2319.
- [18] S. Culurgioni, I.G. Munoz, A. Moreno, A. Palacios, M. Villate, I. Palmero, G. Montoya, F.J. Blanco, The crystal structure of the inhibitor of growth 4 (ING4) dimerization domain reveals the functional organization of the ING family of chromatin binding proteins, *J. Biol. Chem.* 287 (14) (2012) 10876–10884.
- [19] G. Ormazá, B. Medagli, A. Ibanez de Opakua, J.A. Rodríguez, N. Merino, M. Villate, S. Onesti, F.J. Blanco, The tumor suppressor inhibitor of growth 4 binds double-stranded DNA through its disordered central region, *FEBS Lett.* 591 (2) (2017) 425–432.
- [20] V.N. Uversky, Size-exclusion chromatography in structural analysis of intrinsically disordered proteins, *Methods Mol. Biol.* 896 (2012) 179–194.
- [21] K. Dutta, A. Alexandrov, H. Huang, S.M. Pascal, pH-induced folding of an apoptotic coiled coil, *Protein Sci.* 10 (12) (2001) 2531–2540.
- [22] P. Kumar, D.N. Woolfson, Socket2: a program for locating, visualising, and analysing coiled-coil interfaces in protein structures, *Bioinformatics* 37 (23) (2021) 4575–4577.
- [23] P. Gouet, E. Courcelle, D.I. Stuart, F. Metz, ESPrInt: analysis of multiple sequence alignments in PostScript, *Bioinformatics* 15 (4) (1999) 305–308.
- [24] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol.* 372 (3) (2007) 774–797.
- [25] S. Kim, S. Natesan, G. Cornilescu, S. Carlson, M. Tonelli, U.L. McClurg, O. Binda, C. N. Robson, J.L. Markley, S. Balaz, K.C. Glass, Mechanism of histone H3K4me3 recognition by the plant homeodomain of inhibitor of growth 3, *J. Biol. Chem.* 291 (35) (2016) 18326–18341.
- [26] U.L. McClurg, A. Nabbi, C. Ricordel, S. Korolchuk, S. McCracken, R. Heer, L. Wilson, L.M. Butler, B.K. Irving-Hooper, R. Pedoux, C.N. Robson, K.T. Riabowol, O. Binda, Human ex vivo prostate tissue model system identifies ING3 as an oncoprotein, *Br. J. Cancer* 118 (5) (2018) 713–726.
- [27] A. Palacios, F.J. Blanco, Macromolecular crowding increases the affinity of the PHD of ING4 for the histone H3K4me3 mark, *Biomolecules* 10 (2020) 2.
- [28] K. Jaquet, A. Fradet-Turcotte, M. Avvakumov, J.P. Lambert, C. Roques, R. K. Pandita, E. Paquet, P. Herst, A.C. Gingras, T.K. Pandita, G. Legube, Y. Doyon, D. Durocher, J. Cote, The TIP60 core regulates bivalent chromatin recognition by 53BP1 through direct H4K20me binding and H2AK15 acetylation, *Mol. Cell* 62 (3) (2016) 409–421.
- [29] K. Qu, K. Chen, H. Wang, X. Li, Z. Chen, Structure of the NuA4 acetyltransferase complex bound to the nucleosome, *Nature* 610 (7932) (2022) 569–574.
- [30] X. Wang, S. Ahmad, Z. Zhang, J. Cote, G. Cai, Architecture of the Saccharomyces cerevisiae NuA4/TIP60 complex, *Nat. Commun.* 9, (1) (2018) 1147.
- [31] P. Xu, C. Li, Z. Chen, S. Jiang, S. Fan, J. Wang, J. Dai, P. Zhu, Z. Chen, The NuA4 core complex acetylates nucleosomal histone H4 through a double recognition mechanism, *Mol. Cell* 63 (6) (2016) 965–975.

- [32] V.V. Rogov, A. Rozenknop, N.Y. Rogova, F. Lohr, S. Tikole, V. Jaravine, P. Guntert, I. Dikic, V. Dotsch, A universal expression tag for structural and functional studies of proteins, *ChemBiochem* 13 (7) (2012) 959–963.
- [33] F.W. Studier, Protein production by auto-induction in high density shaking cultures, *Protein Expr. Purif.* 41 (1) (2005) 207–234.
- [34] R.C. Tyler, H.K. Sreenath, S. Singh, D.J. Aceti, C.A. Bingman, J.L. Markley, B. G. Fox, Auto-induction medium for the production of [U-15N]- and [U-13C, U-15N]-labeled proteins for NMR screening and structure determination, *Protein Expr. Purif.* 40 (2) (2005) 268–278.
- [35] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: the proteomics server for in-depth protein knowledge and analysis, *Nucleic Acids Res.* 31 (13) (2003) 3784–3788.
- [36] Z. Solyom, M. Schwarten, L. Geist, R. Konrat, D. Willbold, B. Brutscher, BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins, *J. Biomol. NMR* 55 (4) (2013) 311–321.
- [37] K. Kazimierzuk, V.Y. Orekhov, Accelerated NMR spectroscopy by using compressed sensing, *Angew. Chem. Int. Ed. Engl.* 50 (24) (2011) 5556–5559.
- [38] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J. L. Markley, B.D. Sykes, 1H, 13C and 15N chemical shift referencing in biomolecular NMR, *J. Biomol. NMR* 6 (2) (1995) 135–140.
- [39] W. Lee, M. Tonelli, J.L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy, *Bioinformatics* 31 (8) (2015) 1325–1327.
- [40] Y.S. Jung, M. Zweckstetter, Mars — robust automatic backbone assignment of proteins, *J. Biomol. NMR* 30 (1) (2004) 11–23.
- [41] A. Palacios, P. Garcia, D. Padro, E. Lopez-Hernandez, I. Martin, F.J. Blanco, Solution structure and NMR characterization of the binding to methylated histone tails of the plant homeodomain finger of the tumour suppressor ING4, *FEBS Lett.* 580 (30) (2006) 6903–6908.
- [42] W. Kabsch, Xds, *Acta Crystallogr. D Biol. Crystallogr.* 66 (Pt 2) (2010) 125–132.
- [43] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R. J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.* 40 (Pt 4) (2007) 658–674.
- [44] R.M. Keegan, S.J. McNicholas, J.M.H. Thomas, A.J. Simpkin, F. Simkovic, V. Uski, C.C. Ballard, M.D. Winn, K.S. Wilson, D.J. Rigden, Recent developments in MrBUMP: better search-model preparation, graphical interaction with search models, and solution improvement and assessment, *Acta Crystallogr. D Struct. Biol.* 74 (Pt 3) (2018) 167–182.
- [45] D. Liebschner, P.V. Afonine, M.L. Baker, G. Bunkoczi, V.B. Chen, T.I. Croll, B. Hintze, L.W. Hung, S. Jain, A.J. McCoy, N.W. Moriarty, R.D. Oeffner, B.K. Poon, M.G. Prisant, R.J. Read, J.S. Richardson, D.C. Richardson, M.D. Sammito, O. V. Sobolev, D.H. Stockwell, T.C. Terwilliger, A.G. Urzhumtsev, L.L. Videau, C. J. Williams, P.D. Adams, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix, *Acta Crystallogr. D Struct. Biol.* 75 (Pt 10) (2019) 861–877.
- [46] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of coot, *Acta Crystallogr. D Biol. Crystallogr.* 66 (Pt 4) (2010) 486–501.
- [47] V.B. Chen, W.B. Arendall 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G. J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 66 (Pt 1) (2010) 12–21.
- [48] Y. Yan, H. Tao, J. He, S.Y. Huang, The HDock server for integrated protein-protein docking, *Nat. Protoc.* 15 (5) (2020) 1829–1852.



## Trispecific T-cell engagers for dual tumor-targeting of colorectal cancer

Antonio Tapia-Galisteo <sup>a</sup>, Íñigo Sánchez Rodríguez <sup>a</sup>, Oscar Aguilar-Sopeña <sup>b,c,\*</sup>, Seandean Lykke Harwood <sup>d\*</sup>, Javier Narbona <sup>e</sup>, Mariola Ferreras Gutierrez <sup>f</sup>, Rocío Navarro <sup>g</sup>, Laura Martín-García <sup>a</sup>, Cesáreo Corbacho <sup>h</sup>, Marta Compte <sup>g</sup>, Javier Lacadena <sup>e</sup>, Francisco J. Blanco <sup>f</sup>, Patrick Chames <sup>i</sup>, Pedro Roda-Navarro <sup>b,c</sup>, Luis Álvarez-Vallina <sup>j,k</sup>, and Laura Sanz <sup>a</sup>

<sup>a</sup>Molecular Immunology Unit, Biomedical Research Institute Hospital Universitario Puerta de Hierro Majadahonda, Madrid, Spain; <sup>b</sup>Department of Immunology, Ophthalmology and ENT, School of Medicine, Universidad Complutense de Madrid, Spain; <sup>c</sup>Lymphocyte Immunobiology Group, Biomedical Research Institute Hospital 12 de Octubre, Madrid, Spain; <sup>d</sup>Protein Science, Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark; <sup>e</sup>Department of Biochemistry and Molecular Biology, Facultad de Ciencias Químicas, Universidad Complutense de Madrid, Spain; <sup>f</sup>Biomolecular NMR, Centro de Investigaciones Biológicas Margarita Salas-CSIC, Madrid, Spain; <sup>g</sup>Department of Antibody Engineering, Leadartis SI, Madrid, Spain; <sup>h</sup>Pathology Department, Hospital Universitario Puerta de Hierro Majadahonda, Madrid, Spain; <sup>i</sup>Antibody Therapeutics and Immunotargeting Group, Aix Marseille University, CNRS, INSERM, Institute Paoli-Calmettes, CRCM, Marseille, France; <sup>j</sup>Cancer Immunotherapy Unit (UNICA), Hospital Universitario 12 de Octubre, Madrid, Spain; <sup>k</sup>Immuno-oncology and Immunotherapy Group, Biomedical Research Institute Hospital 12 de Octubre, Madrid, Spain

### ABSTRACT

Retargeting of T lymphocytes toward cancer cells by bispecific antibodies has demonstrated its therapeutic potential, with one such antibody approved for the treatment of acute lymphoblastic leukemia (blinatumomab) and several other in clinical trials. However, improvement of their efficacy and selectivity for solid tumors is still required. Here, we describe a novel tandem T-cell recruiting trispecific antibody for the treatment of colorectal cancer (CRC). This construct, termed trispecific T-cell engager (TriTE), consists of a CD3-specific single-chain Fv (scFv) flanked by anti-epidermal growth factor receptor (EGFR) and anti-epithelial cell adhesion molecule (EpCAM) single-domain V<sub>HH</sub> antibodies. The TriTE was well expressed in mammalian and yeast cells, bound the cognate antigens of the three parental antibodies, and enabled the specific cytolysis of EGFR- and/or EpCAM-expressing cancer cells, without inducing T cell activation and cytotoxicity against double-negative (EGFR<sup>-</sup>EpCAM<sup>-</sup>) cancer cells. Bivalent bispecific targeting of double-positive HCT116 cells by TriTE improved *in vitro* potency up to 100-fold compared to single-positive cells and significantly prolonged survival *in vivo*. In addition, it was less efficient at killing single-positive target cells than the corresponding bispecific controls, leading to potentially enhanced tumor specificity. Moreover, dual targeting of two tumor-associated antigens may contribute toward preventing the tumor escape by antigen loss caused by selective pressures from conventional single-targeting T-cell engagers, and may help to overcome antigenic heterogeneity.

### ARTICLE HISTORY

Received 12 August 2021  
Revised 21 January 2022  
Accepted 21 January 2022

### KEYWORDS



Trispecific antibodies; single-domain antibodies; scFv; tandem antibodies; cancer immunotherapy; colorectal cancer

## Introduction


Recombinant DNA technology has allowed the development of a wide variety of multispecific and multivalent antibodies with potentially enhanced anti-tumoral activity and reduced Fc-associated toxicity. At present, single-chain variable fragments (scFv), consisting of VH and VL domains connected by a flexible linker peptide, and the variable domain of heavy-chain only antibodies (V<sub>HH</sub>), are the main building blocks used to generate recombinant Fc-free antibodies.<sup>1</sup> Linking several of these building blocks with different specificities allows the design of bispecific (BsAbs) and trispecific (TsAbs) antibodies that are able to recognize one or two tumor-associated antigens (TAA) and one activating or costimulatory receptor in effector cells, thereby redirecting the immune response specifically toward TAA-expressing cancer cells. In this context, the anti-CD19 x anti-CD3 blinatumomab was the first tandem scFv, known as ‘bispecific T-cell engager’ (BiTE), approved by FDA

(2014) and EMA (2015) for the treatment of B-cell acute lymphoblastic leukemia (B-ALL).<sup>2</sup> Since then, an overwhelming number of BsAbs formats has been developed and dozens of them are currently under evaluation in early phases of clinical trials.<sup>3</sup> BsAb design has evolved from the simple BiTEs to complex platforms such as trimeric antibodies, which allows the generation of trimeric and hexavalent BsAbs.<sup>4</sup>

The next challenge to further enhance the effector functions of immune cells is the generation of trifunctional or trispecific antibodies (TsAb). A recent example is the trifunctional natural killer (NK) cell engager (NKCE), targeting the activating receptors Nkp46 and CD16 on NK cells and a TAA on cancer cells.<sup>5</sup> Similarly, a TsAb that interacts with CD38, CD3 and CD28 enhances both T cell activation and tumor targeting.<sup>6</sup> These two constructs are Fab-based and contain Fc regions, with a molecular weight well above 150 kDa. Interestingly, the CD38 trispecific antibody incorporated a Fc mutation for the

**CONTACT** Laura Sanz  [lsalcober@salud.madrid.org](mailto:lsalcober@salud.madrid.org)  Biomedical Research Institute Hospital Universitario Puerta de Hierro, Joaquín Rodrigo 2, Majadahonda, Madrid, Spain.

\*Íñigo Sánchez Rodríguez, Oscar Aguilar-Sopeña and Seandean Lykke Harwood contributed equally to this work

 Supplemental data for this article can be accessed on the [publisher's website](#).

© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ablation of FcγR binding in order to prevent side effects derived from off-target T cell activation by FcγR-expressing cells.<sup>7</sup>

An alternative for the generation of smaller TsAb formats with improved tumor penetration is the exclusive use of scFv and V<sub>HH</sub> as building blocks. The first checkpoint inhibitory T cell-engaging (CiTE) antibody described consisted of an anti-CD33 scFv fused to an anti-CD3 scFv and the extracellular domain of PD-1 in a single polypeptide chain.<sup>8</sup> Another similar concept was the triplebody, with three scFv fused in tandem recognizing two different TAA and CD16<sup>9–11</sup> or CD3.<sup>12,13</sup> In addition, trispecific killer engagers (TriKE) comprise an anti-CD16 scFv or V<sub>HH</sub> and an anti-TAA scFv crosslinked by the human interleukin-15 moiety.<sup>14–16</sup> Last but not least, the TriTAC format incorporates an anti-albumin VHH for extended serum half-life.<sup>17</sup>

Overall, the great majority of fragment-based TsAb are focused on NK cell activation in order to treat hematological malignancies. Here, we propose a TsAb-based strategy to specifically activate T cells against two different TAA in solid tumor cells, using a format that we have named TriTE (Trispecific T-cell Engager). Dual TAA targeting may provide additional benefits, such as decreasing the risk of immune escape by antigen loss or decreasing on-target off-tumor side effects by improving tumor selectivity. The modularity of our TsAb design allows binding domains to be shuffled, to accommodate the phenotype of different tumors. In this study, we have designed, expressed and characterized an anti-EpCAM x anti-CD3 x anti-EGFR TsAb in the TriTE format.

In antibody-based therapies for colorectal cancer (CRC), EGFR is the most commonly targeted TAA. However, the efficacy of anti-EGFR mAbs cetuximab and panitumumab is limited due to primary and acquired resistance.<sup>18</sup> In addition, high levels of expression of EpCAM have been shown in most of carcinomas and is associated with adhesion, proliferation, migration, and invasion of tumor cells.<sup>19</sup> In fact, the EpCAM x CD3 IgG catumaxomab was the first BsAb approved by EMA (2009), although it was withdrawn in 2017 for commercial reasons. Catumaxomab was administered i.p. since i.v administration was not feasible due to hepatotoxicity, which was attributed to Fc-mediated, off-target T cell activation in the liver.<sup>20</sup>

In this proof-of-concept study, we demonstrate that TriTE antibodies were expressed in a functional state, simultaneously binding to EGFR and EpCAM to improve CD3 clustering on T cells and their cytotoxic effect. Moreover, the TriTE antibody showed an enhanced therapeutic effect *in vivo* compared to that of a control CD3 x EGFR BsAb.

## Results

### Design and expression of a trispecific T-cell engager (TriTE)

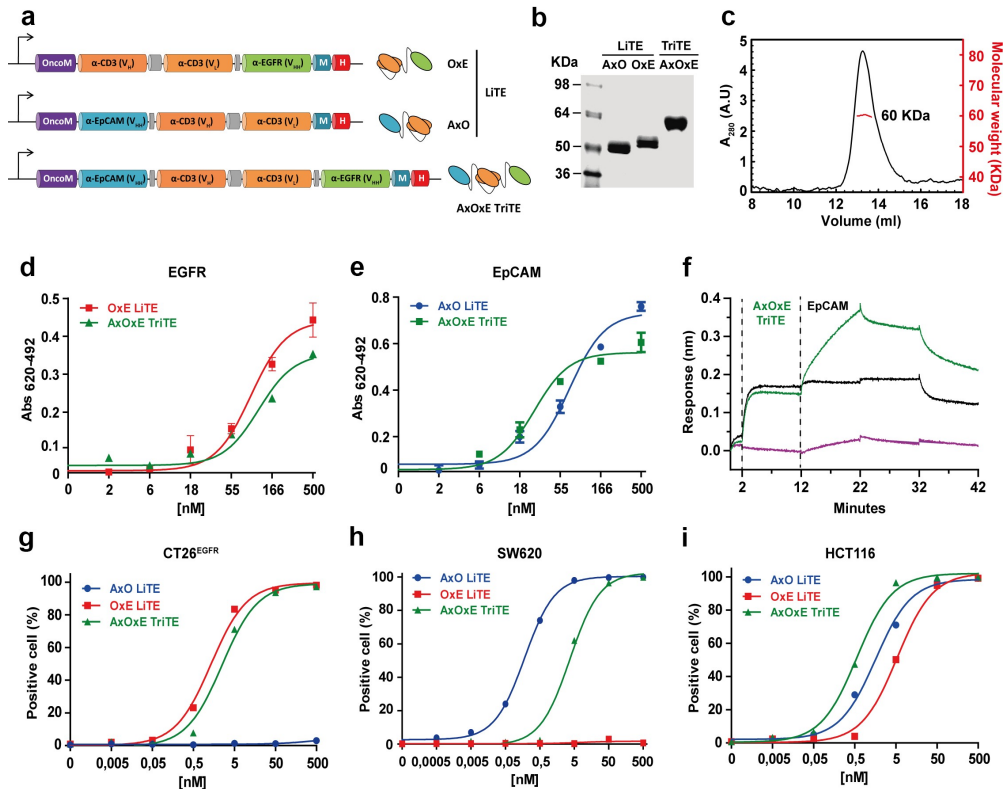
In this study, we generated a trispecific tandem V<sub>HH</sub>-scFv-V<sub>HH</sub> protein (AxOxE) by fusing the anti-human EpCAM A2 V<sub>HH</sub><sup>21</sup> and the anti-human EGFR Ega1 V<sub>HH</sub><sup>22</sup> to the N- and C-terminus, respectively, of the anti-human CD3 OKT3

scFv<sup>23</sup> using flexible five-amino acid (G<sub>4</sub>S) linkers (Figure 1a). This new format was named trispecific T-cell engager (TriTE). For controls, two bispecific light T-cell engagers (LiTEs)<sup>24</sup> were also designed: EpCAM x CD3 (AxO) and CD3 x EGFR (OxE). The three constructs were efficiently produced by transiently transfected human cells. Western blot analysis under reducing conditions of conditioned media (CM) showed a migration pattern consistent with the molecular weights calculated from their amino acid sequence (59 kDa for the AxOxE TriTE, 45 kDa for OxE LiTE and 44 kDa AxO LiTE) and size-exclusion chromatography (SEC) analysis of the AxOxE TriTE in CM showed a peak corresponding to the monomer (Suppl. Figure 1a-b). Next, we demonstrated by ELISA that CM from 293 T<sup>TriTE</sup> and 293 T<sup>LiTE (AxO/OxE)</sup> specifically recognized immobilized EpCAM and/or EGFR (Suppl. Figure 1c). Their ability to detect antigens in a cellular context was also studied by flow cytometry. Fluorescence staining was observed after incubation of CM from 293T<sup>TriTE</sup> and 293T<sup>LiTE OxE</sup> cells with the CRC cell lines HCT116 (EGFR<sup>+</sup>, EpCAM<sup>+</sup>) and CT26<sup>EGFR</sup>, the latter transduced with retrovirus encoding human EGFR<sup>25</sup> (Suppl. Figure 1d). In addition, both AxOxE TriTE and AxO LiTE recognized EpCAM on HCT116 and SW620 cells (EGFR<sup>-</sup>, EpCAM<sup>+</sup>). Furthermore, the three antibodies detected CD3 on the surface of human Jurkat T cells (Suppl. Figure 1d) and were able to activate them in the presence of immobilized target antigens (Suppl. Figure 1e). The monoclonal antibodies cetuximab (anti-EGFR), OKT3 (anti-CD3) and Ber-EP4 (anti-EpCAM) were used as positive controls.

### Purification and characterization of the EpCAMxCD3xEGFR TriTE

For upscaled production, the three antibodies were expressed in *P. pastoris* cells after 24 hours of methanol induction and purified by IMAC, with a yield of roughly 5 mg/L for OxE LiTE and 2 mg/L for AxO LiTE and AxOxE TriTE. Coomassie-stained SDS-PAGE analysis of the purified proteins revealed single bands (>95% pure) with molecular weight slightly higher than the observed in the CM of 293 cells, as previously reported for yeast-produced proteins<sup>26</sup> (Figure 1b, Suppl. Figure 1a). Purified AxOxE TriTE mainly behaved as a globular protein in solution of about 60 kDa as determined by SEC analysis (Figure 1c). Purified AxOxE TriTE and LiTEs showed similar dose-dependent binding curves to plastic-immobilized EGFR or EpCAM, with AxOxE TriTE displaying a slightly lower signal at the highest concentration (Figure 1d-e).

We next studied whether the binding sites of the AxOxE TriTE can bind concurrently to both EGFR and EpCAM. Biolayer interferometry (BLI)-derived sensorgrams showed binding of AxOxE TriTE to EGFR-coated biosensors, giving additional binding curves upon addition of EpCAM (Figure 1f). This experiment demonstrated that the AxOxE TriTE can simultaneously bind to EGFR and EpCAM, and that these interactions are therefore not sterically incompatible.



**Figure 1. Schematic representation and characterization of trispecific T-cell engager (TriTE) and corresponding light T-cell engagers (LiTEs).** (a) Genetic structure of the tandem  $V_{HH}$ -scFv AxO LiTE formed by fusing the anti-EpCAM A2  $V_{HH}$  (blue box) N-terminally to the CD3-specific OKT3 scFv (Orange box); the scFv- $V_{HH}$  OxE LiTE comprising the anti-EGFR Ega1  $V_{HH}$  (green box) fused C-terminally to the OKT3 scFv; and the  $V_{HH}$ -scFv- $V_{HH}$  AxOxE TriTE with anti-EpCAM and anti-EGFR  $V_{HH}$  fused to the N- and C-terminus of OKT3 scFv, respectively. The Oncostatin M signal peptide (purple box) is used to direct secretion of recombinant antibody, and the myc/6xHis tags (dark blue and red boxes) were appended for immunodetection and affinity purification, respectively. Schematic representations showing arrangement of  $V_{HH}$  and scFv in each construct are shown on the right. (b) Reducing SDS-PAGE of the three constructs and (c) SEC analysis of the purified AxOxE TriTE with the indicated molecular weight measured at the center of the chromatography peak (red line). (d-e) Titration ELISA against plastic-immobilized EGFR and EpCAM. Experiments were performed at least twice in duplicates. Mean  $\pm$  SD are shown at each concentration. (f) Biolayer interferometry (BLI)-derived sensorgrams for the interaction between immobilized EGFR and AxOxE TriTE in the presence (green) or not (black) of soluble EpCAM. Note that the TriTE was present during association with EpCAM to prevent dissociation of the TriTE from immobilized EGFR. (g-i) FACS on CT26<sup>EGFR</sup> (EGFR<sup>+</sup>EpCAM<sup>-</sup>), SW620 (EGFR<sup>+</sup>EpCAM<sup>+</sup>) and HCT116 cells (EGFR<sup>+</sup>EpCAM<sup>+</sup>). Percentages of positive cells are shown at each concentration. AxO = A2 (anti-EpCAM  $V_{HH}$ ) x OKT3 (anti-CD3 scFv); OxE = OKT3 x Ega1 (anti-EGFR  $V_{HH}$ ); AxOxE = A2 x OKT3 x Ega1.

Then, the ability of AxOxE TriTE to detect the three cognate antigens as cell surface proteins was studied by flow cytometry (Figure 1g-i, Suppl. Figure 2). All three constructs recognized CD3 on the surface of Jurkat T cells. The AxO LiTE stained EpCAM on the surface of SW620 and HCT116 cells, and the OxE LiTE detected EGFR on CT26<sup>EGFR</sup> and HCT116 cells. Indeed, the AxO LiTE performed better than TriTE on SW620 cells. However, only the AxOxE TriTE was able to stain the three CRC cell lines. The AxO and OxE LiTE did not stain single-positive CT26<sup>EGFR</sup> or SW620 cells, respectively. Neither of the three recombinant antibodies bound to the triple negative CT26<sup>mock</sup> cell line (hereafter referred to as CT26).

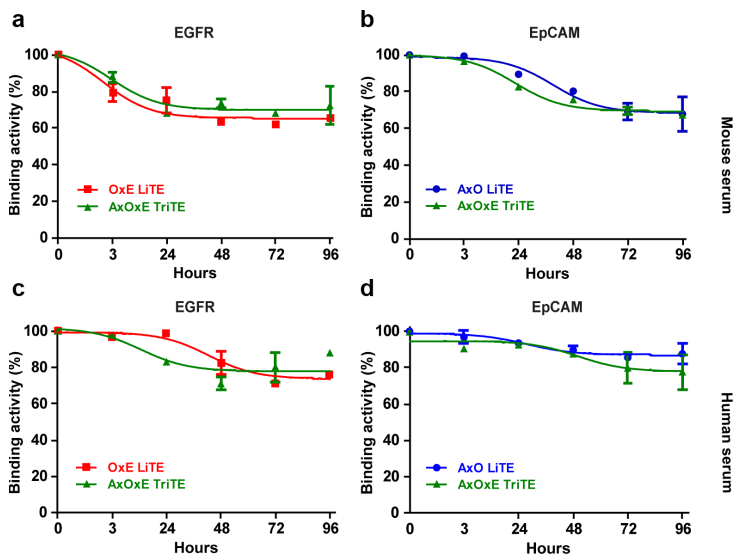
It has been described that apparent affinity increases when a bivalent antibody binds to the second target following its binding to the first receptor on the same cell.<sup>27,28</sup> This phenomenon is also observed in the case of AxOxE TriTE: apparent affinity is 4.5 nM and 3.94 nM for SW620 and CT26<sup>EGFR</sup> cells, respectively, but changes to 0.51 nM in double-positive HCT116 (approximately sevenfold). These differences cannot be

attributed to higher expression of target antigens in HCT116 cells. Indeed, EpCAM expression levels on the surface of SW620 and HCT116 cells are practically identical (MFI = 633 and 640, respectively) (Suppl. Figure 3a). Moreover, CT26<sup>EGFR</sup> cells exhibited slightly higher EGFR levels than HCT116 (MFI = 550 vs 505).

Next, we analyzed the long-term stability in serum of the three constructs. For this purpose, purified proteins were incubated in 60% mouse (Figure 2a-b) or human (Figure 2c-d) serum for 0 (control) to 4 days at 37°C. The purified AxOxE TriTE was very stable with 80% EGFR and EpCAM binding activity after 96 hours of incubation, comparable to those of OxE and AxO LiTEs.

### Inhibition of cell proliferation and EGFR phosphorylation

It has been described that Ega1  $V_{HH}$  is able to block the activity of EGFR by preventing the conformational change of the receptor and thus, its dimerization.<sup>29</sup> To assess the functionality of the



**Figure 2. Serum stability of purified EpCAMxCD3xEGFR TriTE and EpCAMxCD3 and CD3xEGFR LiTEs.** AxOxE TriTE and LiTEs were incubated in mouse (a-b) or human (c-d) serum at 37°C for 96 hours and their functional activity was analyzed by ELISA against plastic-immobilized EGFR (left) or EpCAM (right). Experiments were performed twice in duplicates. Mean  $\pm$  SD are shown at each time point. AxO = A2 (anti-EpCAM V<sub>H</sub>H) x OKT3 (anti-CD3 scFv); OxE = OKT3 x Ega1 (anti-EGFR V<sub>H</sub>H); AxOxE = A2 x OKT3 x Ega1.

Ega1 moiety in the AxOxE TriTE and OxE LiTE, we studied their ability to inhibit proliferation and block EGFR phosphorylation in A431 cells. Notably, the EGFR-dependent A431 cells also express EpCAM.<sup>30</sup> Whereas AxOxE TriTE inhibited A431 proliferation by a 20%, compared with OKT3 ( $P = .005$ ), the OxE LiTE had no significant effect on proliferation, as previously described<sup>31</sup> (Suppl. Figure 4a).

To assess the effect on EGFR phosphorylation status, A431 cells were stimulated with 25 ng/mL of human EGF, after incubation with serial dilutions of AxOxE TriTE or OxE LiTE at the highest concentration. Cetuximab was used as a positive control and untreated cell as negative control. Interestingly, AxOxE TriTE was able to decrease pEGFR in a dose-dependent manner, whereas the OxE LiTE had no effect, in accordance with proliferation results (Suppl. Figure 4b).

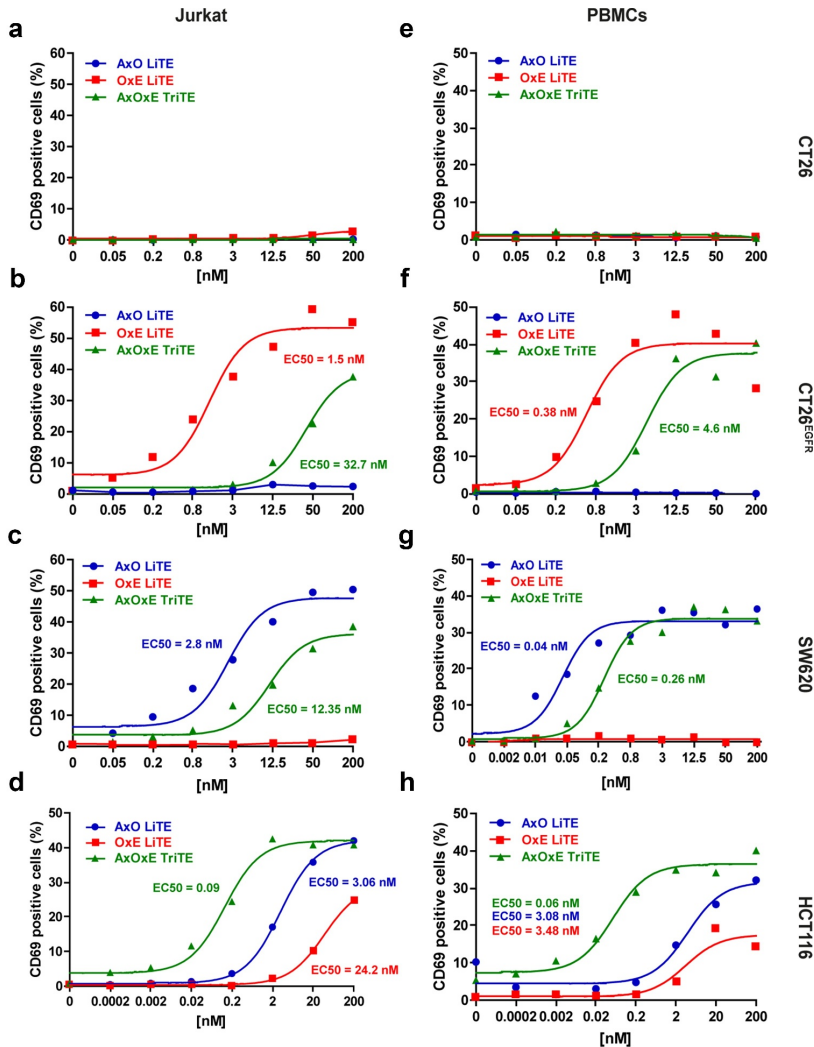
### **EpCAMxCD3xEGFR TriTE specifically activated T cells and triggered the assembly of canonical immunological synapses**

We next assayed the purified AxOxE TriTE and LiTEs for their ability to activate T cells *in vitro*. As shown in Figure 3, LiTEs increased CD69 expression in a dose-dependent manner and more efficiently than AxOxE TriTE when Jurkat cells or PBMCs were co-cultured with single-positive EGFR<sup>+</sup>EpCAM<sup>-</sup> CT26<sup>EGFR</sup> (Figure 3b,f) or EGFR<sup>-</sup>EpCAM<sup>+</sup> SW620 (Figure 3c,g) tumor cells. However, when PBMCs were cocultured with double-positive EGFR<sup>+</sup>EpCAM<sup>+</sup> HCT116 tumor cells in the presence of AxOxE TriTE, the induction of CD69 expression was stronger compared to that

of AxO ( $P = .015$ ) and OxE ( $P = .002$ ) LiTEs (Figure 3d,h). In fact, AxOxE TriTE reached nearly full activation of PBMCs at 2 nM, whereas CD69 expression was almost basal at the same equimolar concentration of both LiTEs. CD69 expression was not induced when T cells were cocultured with double-negative EGFR<sup>-</sup>EpCAM<sup>-</sup> CT26 cells in the presence of AxOxE TriTE or LiTEs (Figure 3a,e).

To check if TriTE activity could be improved by swapping VHH domains, the construct ExOxA was generated. In both TriTE, the VHH in C-ter position was less efficient recognizing single-positive cells or activating T cells cocultured with them (Suppl. Figure 5). However, ExOxA TriTE effect in T cell activation was similar to that of AxOxE in the presence of double-positive HCT116 cells, and superior in both cases to the observed with single-positive cells.

In addition, AxOxE TriTE and LiTEs promoted the formation of the immunological synapse (IS) between Jurkat cells and EGFR<sup>+</sup>EpCAM<sup>+</sup> HCT116 cells as assessed by CD3 $\epsilon$  and F-actin accumulation at the T cell:target cell contact surface (Figure 4a), while the IS was not assembled in the absence of AxOxE TriTE or LiTEs (Suppl. Figure 6). F-actin polarization at the IS was more efficient in the presence of AxOxE TriTE when compared with both LiTEs ( $P < .0001$ ) (Figure 4b), consistent with a higher capability of AxOxE TriTE to activate T cells in comparison with LiTEs (Figure 3d,h). 3D confocal microscopy showed that AxOxE TriTE assembled a canonical IS with peripheral and central distribution of F-actin and CD3 $\epsilon$ , respectively (Figure 4a), whereas this organization was less distinct when LiTEs were used.

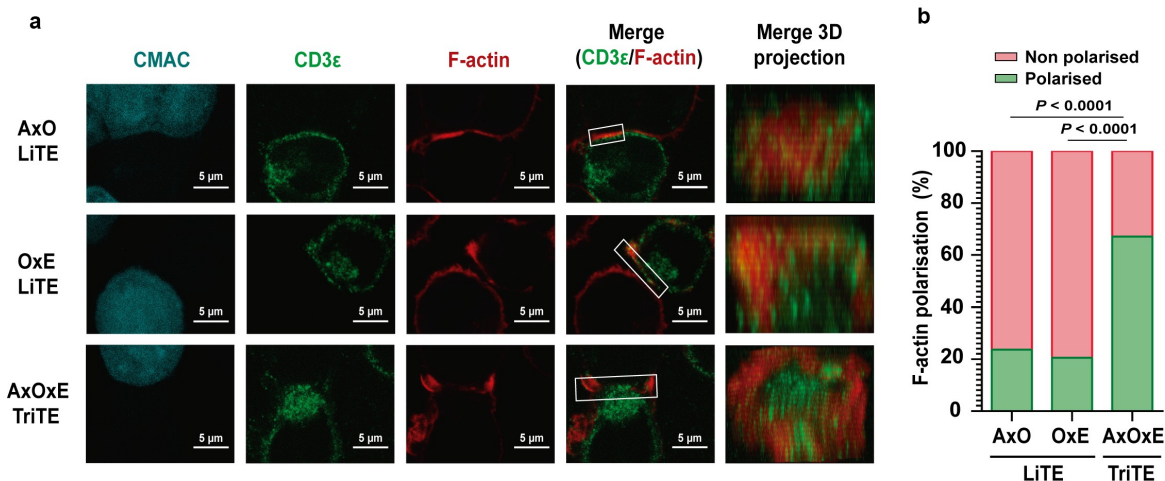


**Figure 3. Induction of T cell activation by purified EpCAMxCD3xEGFR TriTE and EpCAMxCD3 and CD3xEGFR LiTEs.** (a,e) CT26 cells (EGFR<sup>-</sup>EpCAM<sup>+</sup>); (b,f) CT26<sup>EGFR</sup> cells (EGFR<sup>+</sup>EpCAM<sup>+</sup>); (c,g) SW620 cells (EGFR<sup>+</sup>EpCAM<sup>+</sup>) or (d,h) HCT116 cells (EGFR<sup>+</sup>EpCAM<sup>+</sup>) were cocultured with Jurkat cells (left) or PBMCs (right) at the effector/target (E/T) ratio of 5:1 in the presence of different concentrations of purified AxOxE TriTE and LiTEs. After 24 hours, the surface expression of T cell activation marker CD69 was determined by FACS analysis. EC<sub>50</sub> values are provided according to the color code. Experiments were performed three times, one representative experiment is shown. AxO = A2 (anti-EpCAM V<sub>H</sub>H) x OKT3 (anti-CD3 scFv); OxE = OKT3 x Ega1 (anti-EGFR V<sub>H</sub>H); AxOxE = A2 x OKT3 x Ega1.

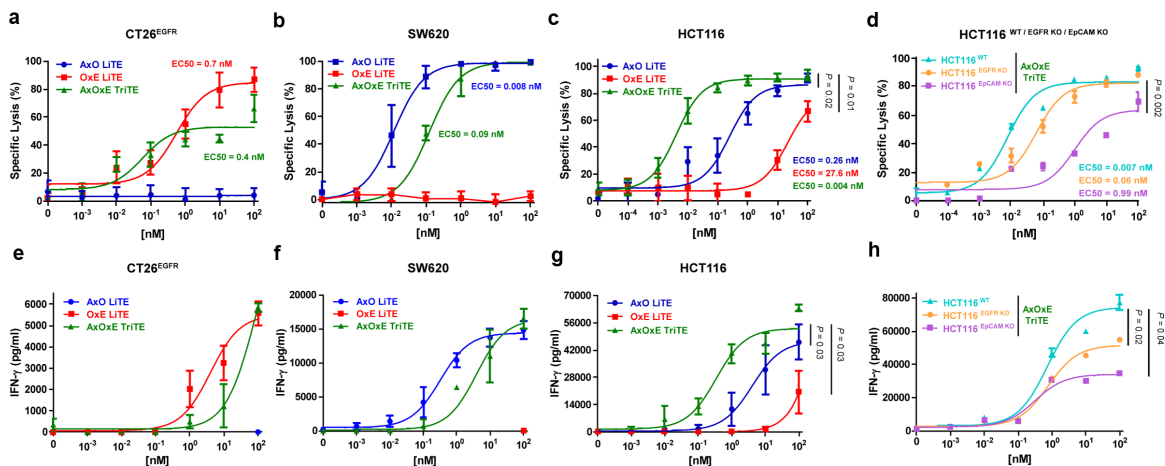
### EpCAMxCD3xEGFR TriTE promoted preferential lysis of EpCAM<sup>+</sup>EGFR<sup>+</sup> cancer cells *in vitro*

We next assessed the ability and specificity of the AxOxE TriTE to elicit cytotoxic responses *in vitro*. PBMCs of three healthy donors were cocultured with CT26<sup>Luc</sup>, CT26<sup>EGFR-Luc</sup>, SW620<sup>Luc</sup> or HCT116<sup>Luc</sup> cells at a 5:1 effector to target (E:T) ratio in the presence of different concentrations of purified antibodies. According to activation data, the three constructs were able to induce dose-dependent killing of EpCAM<sup>+</sup> and/or EGFR<sup>+</sup> cells (Figure 5a-c), which was strictly antigen-specific since it spared EGFR<sup>-</sup>EpCAM<sup>-</sup> CT26 cells (Suppl. Figure 7A). On EGFR<sup>+</sup>EpCAM<sup>+</sup> HCT116<sup>Luc</sup> cells, AxOxE TriTE exhibited a considerably higher cytotoxic ability than

LiTEs, with an EC<sub>50</sub> value of 4 pM, compared with EC<sub>50</sub> values of 0.26 nM and 27 nM for AxO LITE ( $P = .02$ ) and OxE LITE ( $P = .01$ ), respectively (Figure 5c). Moreover, AxOxE TriTE preferentially killed double-positive tumor cells over single-positive ones (EC<sub>50</sub> of 400 pM for CT26<sup>EGFR-Luc</sup> cells and 90 pM in SW620<sup>Luc</sup> cells). This difference could not be attributed to cell line-intrinsic factors other than antigen expression, since cytotoxicity experiments with single antigen-expressing cell lines on the same HCT116 background (KO EGFR<sup>Luc</sup> and KO EpCAM<sup>Luc</sup>) rendered results comparable to those obtained with SW620<sup>Luc</sup> and CT26<sup>EGFR-Luc</sup> cells, respectively, in the presence of AxOxE TriTE (Figure 5d). Expression profiles by FACS of both KO cell lines are shown in Suppl. Figure 3b.



**Figure 4. Immunological synapse formation is triggered by EpCAMxCD3xEGFR TriTE.** (a) Images of immunological synapse (IS) assembly by Jurkat cells co-cultured with EpCAM<sup>+</sup>EGFR<sup>+</sup> HCT116 cells (cyan) in the presence of AxOxE TriTE or LiTEs. The green (CD3ε) and red (F-actin) channels, as well as the merged images, are shown. The IS topology obtained from the 3D reconstructions of regions of interest in confocal stacks (white square) containing the red and the green channels is shown on the right. Experiments were performed three times; results of one representative experiment are shown. Scale bar 5 μm. (b) Percentages of T cells showing F-actin polarization at the IS in each condition are shown. Statistical differences were examined by two-coiled chi-square test. AxO = A2 (anti-EpCAM V<sub>HH</sub>) x OKT3 (anti-CD3 scFv); OxE = OKT3 x Ega1 (anti-EGFR V<sub>HH</sub>); AxOxE = A2 x OKT3 x Ega1.



**Figure 5. Specific cytotoxicity and IFN-γ secretion elicited by purified EpCAMxCD3xEGFR TriTE and EpCAMxCD3 and CD3xEGFR LiTEs.** CT26<sup>EGFR-Luc</sup> cells (a,e), SW620<sup>Luc</sup> cells (b,f) or HCT116<sup>Luc</sup> cells (c,g) were cocultured in 96-well plates with PBMCs at the effector/target (E/T) ratio of 5:1 in the presence of different concentrations of purified AxOxE TriTE and LiTEs. In additional experiments, HCT116<sup>Luc</sup> were compared with the corresponding EGFR and EpCAM knockout cell lines in the presence of TriTE serial dilutions (d,h). After 72 hours, specific cytolysis of tumor cells were measured by bioluminescence assay (upper) and IFN-γ production was determined in CM by ELISA (lower). Percent specific lysis was calculated relative to an equal number of tumor cells cultured with PBMCs in the absence of purified antibodies. EC50 values are provided according to the color code. PBMCs were obtained from 3 different donors, and experiments were performed in triplicate. Statistical differences were examined by unpaired Student's t-test assuming a normal distribution. Results are expressed as a mean ± SD. AxO = A2 (anti-EpCAM V<sub>HH</sub>) x OKT3 (anti-CD3 scFv); OxE = OKT3 x Ega1 (anti-EGFR V<sub>HH</sub>); AxOxE = A2 x OKT3 x Ega1.

**EpCAMxCD3xEGFR TriTE induced strong IFN-γ secretion in vitro**

Next, we analyzed IFN-γ secretion by PBMCs that were cocultured 72 h with tumor cells in the presence of the three antibodies. IFN-γ secretion elicited by AxOxE TriTE was

significantly higher than that of AxO and OxE LiTEs ( $P = .03$  and  $P = .036$ , respectively) in the cocultures with HCT116<sup>Luc</sup> cells (Figure 5g), whereas there were no statistically significant differences between AxOxE TriTE and LiTEs in the induction of IFN-γ in the cocultures with CT26<sup>EGFR-Luc</sup> (Figure 5e) or

SW620<sup>Luc</sup> (Figure 5f). In accordance with cytotoxicity results, AxOxE TriTE promoted higher IFN- $\gamma$  secretion in cocultures with HCT116 wild type than in the presence of HCT116 EGFR or EpCAM KO (Figure 5h).

Importantly, in the presence of CT26 target cells, analyses of CM revealed no increase of IFN- $\gamma$  secretion even at 100 nM, the highest concentration of antibodies used (Suppl. Figure 7B).

### Antitumor effect of EpCAMxCD3xEGFR TriTE

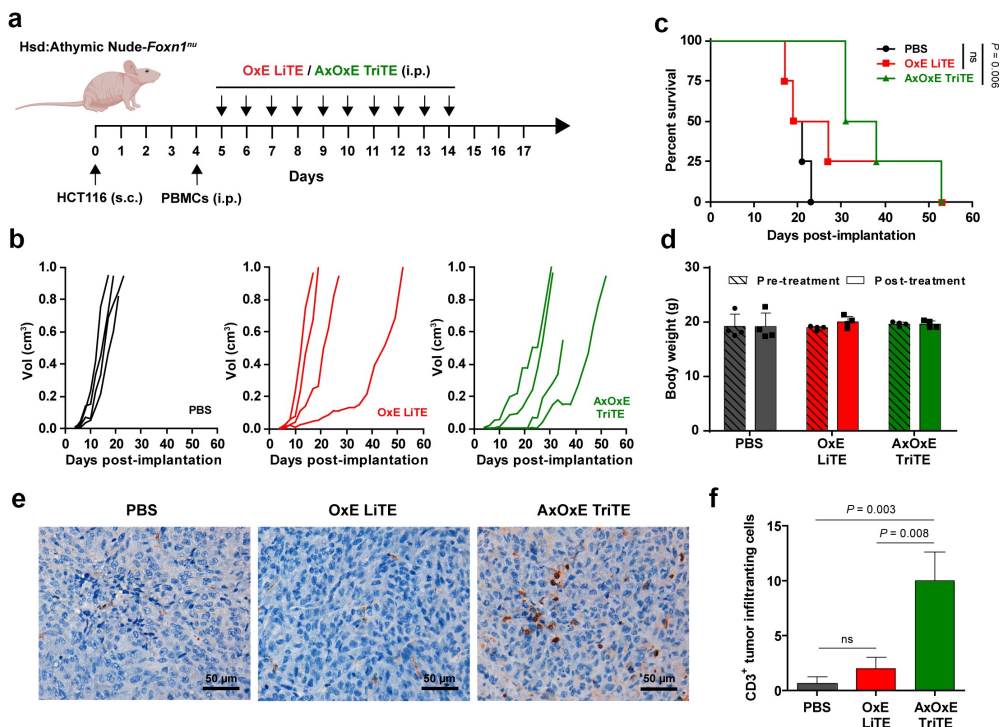
To study the therapeutic effect of AxOxE TriTE *in vivo*, HCT116 cells were implanted in Hsd:Athymic Nude-Foxn1<sup>nu</sup> mice. When the tumors reached average diameters of 0.2 cm (day 4), mice were randomized and human PBMCs were administered intraperitoneally. On day 5, treatment with OxE LiTE or AxOxE TriTE was initiated. Equimolar doses of the antibodies (3 mg/kg for OxE LiTE and 4 mg/kg for AxOxE TriTE) were administered i.p. daily for 10 days (Figure 6a). As shown in Figure 6b, HCT116 tumor-bearing mice treated with AxOxE TriTE showed a delay in tumor growth when compared to their counterparts treated with OxE LiTE. Moreover, 2 out of 4 mice in the TriTE group controlled tumor growth at least until day 21 (Suppl. Figure 8). By day 27 after HCT116 inoculation, all control mice treated with PBS had

been euthanized, whereas 100% of mice in the AxOxE TriTE group were alive, vs. 25% in the group receiving OxE LiTE. In fact, only AxOxE TriTE-treated mice showed a statistically significant increased survival in comparison to the PBS group ( $P = .006$ ) (Figure 6c). No significant differences in body weight were observed before and after the treatment (Figure 6d).

In order to assess the ability of the antibodies to promote immune infiltration, resected tumors were analyzed by immunohistochemistry. Numbers of infiltrating CD3<sup>+</sup> T cells were significantly higher in AxOxE TriTE-treated mice than in PBS- ( $P = .003$ ) or OxE LiTE-treated groups ( $P = .008$ ) (Figure 6e-f).

### Discussion

Intratumor heterogeneity has been associated with poor outcome and decreased response to therapy in a variety of human cancer types, suggesting a universal role in therapeutic resistance. Preexisting heterogeneity increases the risk of at least some tumor cells surviving therapy-induced elimination, while ongoing diversification of tumor cell phenotypes during treatment enables tumor cells to adapt to therapeutic selective pressure, leading to *de novo* resistance.<sup>32</sup> For example, carcinoembryonic antigen (CEA) expression



**Figure 6.** *In vivo* therapeutic effect of EpCAMxCD3xEGFR TriTE and CD3xEGFR LiTE. (a) Hsd:Athymic Nude-Foxn1<sup>nu</sup> mice were inoculated subcutaneously with  $2 \times 10^6$  HCT116 tumor cells. Mice were randomized into groups ( $n = 4$ /group) when tumors reached 0.2 cm in diameter and injected intraperitoneally with  $10^7$  freshly isolated PBMCs. Then, mice were treated with intraperitoneal injections of PBS, OxE LiTE or AxOxE TriTE daily. (b) Tumor volume growth curves for individual mice are represented. (c) Kaplan-Meier survival curves of AxOxE TriTE- and OxE LiTE-treated mice, log-rank (Mantel-Cox) test. (d) Mice weight before and after treatment are shown. (e) Representative images of CD3<sup>+</sup> TIL immunostaining in tissue sections from HCT116 tumors treated with PBS, OxE LiTE or AxOxE TriTE. Tumors were resected at termination of the experiment shown above. Error bar = 50  $\mu$ m. (f) Quantification of CD3<sup>+</sup> TILs in three independent fields/tissue section. Statistical differences were examined by unpaired Student's t-test assuming a normal distribution. Results are expressed as a mean  $\pm$  SD. OxE = OKT3 (anti-CD3 scFv) x Ega1 (anti-EGFR V<sub>HH</sub>), AxOxE = A2 (anti-EpCAM V<sub>HH</sub>) x OKT3 x Ega1.

heterogeneity and plasticity contribute to resistance to the anti-CEA T-cell bispecific antibody cibisatamab (CEA-CD3-TCB) in patient-derived CRC organoids through CEA antigen loss.<sup>33</sup> In patients treated with CD19-directed immunotherapies, such as chimeric antigen receptor (CAR)-engineered T-cells (CAR-T) and the BiTE blinatumomab, up to 60% and 30% of relapses, respectively, are associated with the loss of CD19, rendering the malignant cells invisible to CD19-specific immunotherapies.<sup>34</sup>

Another escape strategy is the acquisition of mutations that prevent the recognition by the targeting domain, as observed in CRC patients treated with the anti-EGFR mAb cetuximab. Emergence of mutations in the EGFR ectodomain, located in the region of interaction with cetuximab, may disturb this interface and confer resistance to the treatment.<sup>35,36</sup> Interestingly, a subset of *EGFR* mutations preventing binding to cetuximab are still permissive for interaction with panitumumab, the second anti-EGFR mAb approved for CRC treatment.<sup>37,38</sup>

While effective, single-targeted T-cell engagers such as BiTE and TCB exert selective pressure against a unique target antigen which may lead to tumor relapse. Therefore, simultaneous dual targeted T-cell-redirecting therapies may improve tumor specificity while limiting the risk of immune escape.<sup>39</sup> For example, T cells expressing both anti-CD19 and anti-CD123 CARs simultaneously provided superior *in vivo* activity against B-ALL compared with single-expressing CAR-T-cells.<sup>34</sup> Recently, a first-in-human trial of bispecific anti-CD20, anti-CD19 CAR-T-cells for relapsed, refractory B cell malignancies has been reported,<sup>40</sup> demonstrating their potential to overcome the antigen loss observed for single-targeted CD19 CAR-T cells.

Despite promising results with dual targeted CAR-T-cells, a similar approach has not been pursued with T-cell redirecting antibodies. Here, we present a novel trisppecific T-cell engager (TriTE) as an evolution of the BiTE and LiTE formats. This TriTE was generated by the fusion of an anti-CD3 scFv to single-domain anti-EpCAM and anti-EGFR  $V_{HH}$ , for dual targeting of CRC cells. The AxOxE TriTE was secreted as soluble and functional protein by transfected mammalian and yeast cells, recognized the three cognate antigens, and selectively activated and recruited cytolytic human T cells to kill EpCAM<sup>+</sup> and/or EGFR<sup>+</sup> cancer cells *in vitro*. As  $V_{HH}$  lack the hydrophobic interface mediating interaction with VL,<sup>41,42</sup> chain mispairing within the TriTE can be ruled out. Importantly, AxOxE TriTE had no effect when human T cells were cultured with double-negative EpCAM<sup>-</sup>EGFR<sup>-</sup> cells.

Another issue to take into account is that most antigens targeted by therapeutic antibodies are tumor-associated, but not unique to tumor cells: that is to say, they are also expressed at lower levels in normal tissues. EpCAM expression on healthy epithelia of the gastrointestinal tract has limited the therapeutic window of EpCAM-directed therapies due to on-target/off tumor side effects.<sup>19</sup> The mAb edrecolomab, the first anti-EpCAM mAb approved for the treatment of CRC, was subsequently withdrawn when larger studies showed no benefit compared with standard chemotherapy.<sup>43</sup> While edrecolomab efficacy may have been impaired by low binding affinity, the mAbs ING-1 and 3622W94 displayed such a high affinity that they no longer discriminated between normal and malignant cells, and risk of pancreatitis precluded further studies as monotherapy.<sup>44</sup>

A moderate binding affinity could account for the larger therapeutic window observed in patients treated with the anti-EpCAM mAb adecatumumab.<sup>45</sup>

In order to avoid systemic toxicity, it has been proposed that a dual-targeted anti-tumor BsAb should preferentially bind to malignant cells rather than normal cells if the affinity of the individual binding domains are sufficiently low as to require the presence of both target antigens for efficient binding (through the avidity effect),<sup>46</sup> whereas high-affinity binding domains may efficiently bind normal cells expressing only a single antigen and thereby induce off-tumor toxicity. In some circumstances, dual targeting alone may not be sufficient to guarantee selective tumor-targeting, and affinity fine-tuning of one of the binding domains may be required.<sup>47</sup> Affinity-reduced BsAb variants have been shown to mediate a greater degree of tumor selectivity, while the overall therapeutic effect was not ameliorated.<sup>48,49</sup> Interestingly, AxOxE TriTE is less efficient in activating T cells (as assessed by CD69 expression and IFN- $\gamma$  secretion) and triggering their cytotoxic effect on single positive cells than the corresponding LiTEs. This may be a consequence of its engineering in a single-chain polypeptide that may partially interfere at certain extent with antigen recognition by individual binding domains. In fact, AxOxE TriTE  $EC_{50}$  is 100 and 22.5 times higher in CT26<sup>EGFR</sup> and SW620 than in HCT116, respectively. The anti-EGFR  $V_{HH}$  seems to be especially impaired in the C-terminal position of AxOxE TriTE, since its capacity to kill CT26<sup>EGFR</sup> cells is considerably reduced in comparison with the corresponding anti-EGFR LiTE. The positioning effect on binding affinity of anti-EGFR  $V_{HH}$  placed in C-terminal end has been previously described.<sup>17</sup> Although this reduction was not deliberate in our case, we can speculate that it may render TriTE able to discriminate to some extent between double-positive and single positive cells, or cells with high vs. low antigen density. Indeed, AxOxE TriTE performs much better in cytotoxicity assays with double-positive HCT116 cells than LiTEs, with  $EC_{50}$  65 and 675 times lower than that of AxO LiTE and OxE LiTEs, respectively, suggesting a potentially more favorable safety profile, while preserving its therapeutic effect *in vivo*. Discrimination could be further enhanced generating AxOxE TriTE variants through alanine mutagenesis of targeted residues in the CDR3 of anti-EpCAM binding domain as previously described,<sup>49</sup> in order to completely abrogate binding to single positive cells, while preserving recognition of double-positive ones. Affinity-reduced variants could also distinguish between high antigen density malignant cells and normal cells with low antigen density.

Interestingly, AxOxE TriTE partially recapitulated the effect of bivalent cetuximab in the inhibition of EGFR-mediated signaling. It had been previously reported that monovalent binding to EGFR by a LiTE had no effect on proliferation or receptor phosphorylation status of A431 cells.<sup>31</sup> Although EGFR binding by AxOxE TriTE is also monovalent, this difference may be attributed to EpCAM expression in A431 cells<sup>30</sup> and thus the higher avidity of AxOxE TriTE. This could theoretically allow the AxOxE TriTE to function as a dual mechanism therapeutic,



redirecting T cells toward CRC cells while simultaneously inhibiting mitogenic signaling from EGFR in the CRC cells.

In summary, we demonstrate the potential of the tandem trispecific T-cell engager format, using anti-EpCAM and anti-EGFR binding domains for proof of concept. Moreover, a wide selection of TAA-specific  $V_{HH}$  antibodies are available, and their combination with well-characterized anti-CD3 or anti-CD16 binding domains could easily provide a therapeutic arsenal of TriTEs and TriKEs aimed toward treatment of diverse cancers, with potentially enhanced efficacy, increased tumor selectivity and reduced risk of clonal escape.

## Materials and methods

### General reagents and antibodies

The human EGFR-Fc (cat#344-ER) and EpCAM-Fc (cat#960-EP) chimeras were from R&D Systems (Minneapolis, MN) and bovine serum albumin (cat#A9647, BSA) was from Sigma-Aldrich (St. Louis, MO). The mAbs used included: mouse anti-c-myc clone 9E10 (cat#ab206486, Abcam, Cambridge, UK), mouse anti-human CD3 $\epsilon$  clone OKT3 (Ortho Biotech, Bridgewater, NJ), chimeric anti-human epidermal growth factor receptor (EGFR) cetuximab (Merck KGaA, Darmstadt, Germany), mouse anti-human epithelial antigen clone Ber-EP4 (cat#M0804, Dako), phycoerythrin (PE)-conjugated anti-human CD69 clone FN50 (cat#555531, BD Biosciences, San Jose, CA), fluorescein isothiocyanate (FITC)-conjugated anti-human CD3 clone OKT3 (cat#566783, BD Biosciences), the rabbit anti-human phospho-EGFR (Tyr1068) clone D7A5 (cat#3777, Cell Signaling Technology Inc) and anti- $\beta$ -actin mouse clone 8226 (cat#ab8226, Abcam). The polyclonal antibodies included: PE-conjugated goat F(ab')<sub>2</sub> fragment anti-mouse IgG, Fc specific, (cat#115-116-071, Jackson Immuno Research, Newmarket, UK); PE-conjugated goat F(ab')<sub>2</sub> fragment anti-human IgG (H&L) (cat#109-116-170, Jackson Immuno Research, Newmarket, UK), horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG (cat#A5278, Sigma-Aldrich), HRP-conjugated goat anti-human IgG (cat#A-0170, Sigma-Aldrich), IRDye800CW-donkey anti-rabbit (cat#925-32213, LI-COR Biosciences) and IRDye680RD-donkey anti-mouse (cat#925-68072, LI-COR Biosciences).

The following reagents were also used: human EGF (cat#130-093-825, Miltenyi Biotec), D-luciferine (cat#E160C, Promega, Madison, WI, USA), CellTiter-Glo luminescent assay (cat#G7571, Promega), glycerol (cat#56-81-5, VWR LifeScience), methanol (cat#K977, Amresco), yeast extract (cat#1702.00, Condalab), peptone (cat#1616.00, Condalab), YNB (cat#1545.00, Condalab) and dextrose (cat#X997.2, Roth). Cetuximab was obtained from the pharmacy at Hospital Puerta de Hierro.

### Cells and culture conditions

Human 293T (CRL-3216), SW620 (CCL-227), HCT116 (CCL-247) and Jurkat (TIB-152) cells were obtained from ATCC (American Type Culture Collection, Rockville, MD, USA). Mouse CT26 cells (CRL-2638) infected with p-BABE-

puro-EGFR expressing human EGFR (CT26<sup>EGFR</sup>) or infected with the empty vector retrovirus (CT26<sup>mock</sup>) were provided by M. Rescigno.<sup>25</sup> EGFR knockout (ab281597) and EpCAM knockout (ab281596) HCT116 cell lines were purchased from Abcam. Production of lentiviral vectors for the generation of HCT116 (WT and KO), SW620, CT26<sup>mock</sup> and CT26<sup>EGFR</sup> cells expressing the firefly luciferase (Luc) gene has been described previously.<sup>50</sup> The epidermoid carcinoma cell-line A431, carrying an amplification of the EGFR gene, was obtained from the ATCC (CRL-1555) and FreeStyle<sup>TM</sup> 293 F cells were provided by Invitrogen (R790-07). PBMCs from healthy donors were provided by the Biobank of Hospital Universitario Puerta de Hierro Majadahonda (HUPHM)/Instituto de Investigación Sanitaria Puerta de Hierro-Segovia de Arana (IDIPHISA) (PT17/0015/0020, Spanish National Biobank Network), with the appropriate approval of the Ethics Committee and based on informed consent. Adherent cells were cultured in DMEM medium (Lonza, Walkersville, MD) supplemented with 10% FCS (Sigma-Aldrich, St. Louis, MO) and 1% pen-strep-glutamine (Gibco, Thermo Fisher Scientific, Waltham, MA). Jurkat cells and PBMCs were maintained in RPMI-1640 (Lonza) supplemented with 10% FCS and 1% pen-strep-glutamine. FreeStyle<sup>TM</sup> 293 F cells were cultured in FreeStyle<sup>TM</sup> 293 expression medium (Invitrogen). All cells were routinely screened for mycoplasma contamination by PCR (Biotools, Madrid, Spain) at the Tissue Culture Core Facility, Biomedical Research Institute Puerta de Hierro-Segovia de Arana and were authenticated at the Universidad Complutense de Madrid Genomics Unit using the AmpFLSTR Identifier PCR Amplification kit (Applied Biosystems, Thermo Fisher Scientific). *Pichia pastoris* KM71H strain was provided by Dr. Javier Lacadena Gallo (UCM) and cells were cultured with YPD (1% yeast extract, 2% peptone, 2% dextrose), BMXY (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer pH 6.0,  $4 \times 10^{-5}$  biotin) supplemented with 1% glycerol (BMGY) or 0.5% methanol (BMMY).

### Construction of expression vectors

To generate the EpCAMxCD3 (AxO) and CD3xEGFR (OxE) LiTEs expression vectors, DNA fragments encoding the anti-EpCAM A2 (sequence kindly provided by Patrick Chames) and anti-EGFR Ega1  $V_{HH}$ <sup>29</sup> were synthesized by Genart AG (Thermo Fisher Scientific), digested with ClaI/NotI or XhoI/EcoRI, respectively, and ligated into the pCR3.1-OKT3 plasmid. To generate the AxOxE TriTE construct, the PCR fragment Ega1 cleaved with XhoI/EcoRI was ligated into a pCR3.1 vector containing the coding sequence of OKT3, and A2  $V_{HH}$  was cloned via ClaI/NotI, resulting in pCR3.1- AxOxE TriTE expression vector. Finally, in order to obtain the ExOxA construct, PCR fragments Ega1 and A2 were digested with ClaI/NotI or XhoI/EcoRI, respectively, and ligated into the same vector. Individual binding domains in each construct were connected by short linkers (G<sub>4</sub>S). For medium-scale protein production, the three constructs were subcloned into the ClaI/XbaI digested pPicza vector (Invitrogen), to obtain pPicza-AxO, pPicza-OxE LiTEs and pPicza-AxOxE TriTE. All plasmids were

amplified in chemically competent *Escherichia coli* TOP10 and purified using Qiagen plasmid Midi kit. Final sequences were verified using F-CMV and R-BGH or 5'-AOX1 and 3'-AOX1.

### Expression in mammalian cells

293T cells were transiently transfected with pCR3.1 vectors encoding the three antibodies using calcium phosphate and CM were collected after 48 h. FreeStyle<sup>TM</sup> 293 F cells ( $10 \times 10^7$ ) were transfected with a ratio 1:1 of pCR3.1 – AxOxE TriTE (100 µg) and PEI (100 µg) in 100 ml of FreeStyle<sup>TM</sup> expression medium. Antibody expression was analyzed using ELISA and Western blotting.

### Expression in *P. pastoris* and purification of recombinant antibodies

Electrocompetent *P. pastoris* KM71H strain cells were electroporated with 10 µg of appropriate linearized pPicza plasmids after digestion with PmeI, using Bio-Rad Gene pulser apparatus (Bio-Rad, Hercules, CA, USA). The yeast cells containing the integrated sequences were selected in YPD plates with 100, 400 or 750 µg/mL of zeocin. Different clones were tested to select those with the highest yield. For this purpose, individual clones were grown in BMGY (buffered media for yeast containing glycerol) for 24 hours at 30°C and 200 rpm and protein production was induced with BMMY (buffered media for yeast containing methanol) at 15°C and 200 rpm. Methanol 0,5% (v/v) was added every 24 hours and samples of CM were collected to analyze the expression profile by SDS-PAGE and Western blot. Colonies with the highest production of each construction were selected for medium-scale production, grown in baffled erlenmeyer flasks with 2–4 L of BMGY medium at 30°C, 200 rpm for 24 hours and induced with 200–400 mL of BMMY medium at 15°C, 200 rpm for 24 hours. Collected cell-free CM were dialyzed against PBS 1x at 4°C for 24 hours and recombinant antibodies were purified by affinity chromatography with HisTrap<sup>TM</sup> HP columns (GE Healthcare) using an ÄKTA Prime plus system (GE Healthcare, Uppsala, Sweden). Endotoxin levels were <0.25 EU/ml as determined by the LAL Endotoxin Kit (Pierce).

### Size exclusion chromatography

A sample of 200 µL of CM from transfected 293 cells was injected into a Superdex 200 Increase 10/300 GL column (Cytiva, MA, US) on an ÄKTA GO chromatography system (Cytiva) at room temperature, while monitoring light absorbance at 280 nm. The column was equilibrated in phosphate buffered saline pH 7.4 plus 150 mM NaCl and run in the same buffer at 0.5 mL per minute. The column was previously calibrated with a set of Gel Filtration Standards (Biorad, from 1.4 to 670 kDa). The fractions containing monomeric protein were concentrated and reinjected under the same conditions to assess that the monomeric protein was stable and did not aggregate. To check the molecular size of purified AxOxE

TriTE in solution, SEC was also performed. An aliquot corresponding to the elution volume of the AxOxE TriTE was rechromatographed to assess its monomeric state.

### Western blot

Samples of cell-free CM or purified proteins were analyzed under reducing conditions on 12% Tris-glycine gels and transferred to nitrocellulose membranes using iBlot system (Life Technologies). After incubation with LI-COR blocking solution (LI-COR, Lincoln, NE, USA), proteins were detected with 1 µg/mL mouse anti-c-myc mAb (9E10, cat#ab206486, Abcam), followed by DyLight800-conjugated goat anti-mouse IgG (cat#610-145-121, Rockland Immunochemicals) diluted 1:5000. Visualization of protein bands was performed with the Odyssey system (LI-COR Biosciences).

### Enzyme-linked immunosorbent assay

Human EGFR-Fc or huEpCAM-Fc chimeras were immobilized (5 µg/mL) on Maxisorp plates (NUNC Brand Products) overnight at 4°C. After washing and blocking with Odyssey blocking buffer (Li-Cor Biosciences), CM or increasing amounts of purified antibodies were added for 1 hour at room temperature. The wells were washed and incubated with 1 µg/mL anti-c-myc mAb for 1 hour at room temperature. After washing, proteins were detected with HRP-conjugated goat-anti-mouse IgG (1:1000 dilution) (cat#A5278, Sigma-Aldrich) for 45 minutes at room temperature. Finally, the plates were developed using O-phenylenediamine dihydrochloride (OPD) in citrate phosphate buffer and the reaction was stopped using sulfuric acid 1 M. Cetuximab (5 µg/mL) (pharmacy at Hospital Puerta de Hierro) and Ber-EP4 (1:200) were used as positive controls and detected with HRP-conjugated goat-anti-human IgG (1:1000 dilution) or HRP-conjugated goat-anti-mouse IgG (1:1000 dilution), respectively.

### Biolayer interferometry

The simultaneous binding of AxOxE TriTE to immobilized huEGFR-Fc (R&D Systems) and huEpCAM-Fc (R&D Systems) in solution was measured using biolayer interferometry on an Octet RED96 system (Fortebio). Prior to the experiment, anti-hFc capture biosensors (Fortebio) were incubated with 40 nM of huEGFR-Fc for 20 minutes in HEPES-buffered saline (HBS, 20 mM HEPES, 150 mM NaCl, pH 7.4). The antibody was loaded onto the immobilized huEGFR-Fc at 100 nM for 10 minutes in HBS. One biosensor was then moved into a solution still containing 100 nM of the AxOxE TriTE, and another into a solution with both 100 nM of the antibody and 300 nM of EpCAM. A control biosensor which was loaded with huEGFR-Fc but not the AxOxE TriTE also was also incubated with 300 nM of EpCAM. After monitoring the association with EpCAM for 10 minutes, the biosensors were moved back into solutions containing only 100 nM of the antibody to measure EpCAM dissociation for 10 minutes, and then moved into HBS only to monitor antibody dissociation for an additional 10 minutes.

### Flow cytometry

CT26, CT26<sup>EGFR</sup>, SW620, HCT116 or Jurkat cells were incubated with CM or purified antibodies for 1 hour on ice. In case of titration experiments, the antibodies were used in tenfold dilution series spanning the concentration range from 500 nM down to 0.5 pM. After washing, 1 µg/mL anti-c-myc mAb was added for 1 hour at 4°C and detected using a phycoerythrin (PE)-conjugated goat-anti-mouse antibody (1:200 dilution). Cetuximab, Ber-EP4 and OKT3 mAbs were used as positive controls. Cells incubated without primary antibody were used as negative controls. Samples were acquired on a MACSQuant Analyzer 10 (Miltenyi Biotec, Bergisch Gladbach, Germany) and analyzed using FlowJo (BD Biosciences, Franklin Lakes, NJ, USA) at the Flow Cytometry Core Facility, Biomedical Research Institute Puerta de Hierro-Segovia de Arana.

### Serum stability

Purified antibodies (6 µg) were incubated in PBS 60% human and mouse serum at 37°C for 5 days. Samples were collected at 3 and every 24 hours and their binding activities were tested by ELISA, representing the sample at 0 hours 100% of functionality.

### T cell Activation assays

Microtiter 96-well plates were seeded with tumor cells ( $2 \times 10^4$ /well) 24 hours before. Consecutively, wells were incubated with CM or increasing amounts of purified antibodies for 30 min at 37°C. After washing, Jurkat cells or human PBMCs, isolated from healthy volunteers by density-gradient centrifugation, were added at 5:1 effector:target (E:T) ratio. After 24 hours, the expression profile of activation marker CD69 was determined by FACS using a PE-conjugated anti-CD69 mAb and FITC-conjugated anti-CD3 mAb incubated for 30 minutes on ice. Samples were analyzed with a MACSQuant Analyzer 10 (Miltenyi Biotec GmbH).

### Immunological synapse formation

HCT116 cells were labeled with 1 µM cell tracker dye 7-amino-4-chloromethylcoumarin (CMAC; Life Technologies) and incubated with 5 nM AxOxE TriTE or LiTEs for 30 minutes. After washing, an equal number of Jurkat and HCT116 cells were co-incubated for 15 minutes on poly-L-lysine-coated coverslips at 37°C in a humidified atmosphere with 5% CO<sub>2</sub>. Samples were fixed with 4% paraformaldehyde for 5 minutes at room temperature and permeabilised with TBS-Triton 0.1% for 5 minutes at room temperature. After blocking for 20 minutes with 10 µg/ml human gamma globulin (Sigma-Aldrich), samples were stained with mouse α-human CD3ε antiserum (kindly provided by Francisco Sanchez-Madrid, Hospital Universitario de la Princesa, Madrid, Spain) diluted 1/2 in TNB buffer (Roche Diagnostics) for 1 hour at room temperature. After washing with TBS, cells were incubated with Alexa Fluor 488-conjugated donkey anti-rabbit antibody (1:500; Life Technologies) and phalloidin-Alexa-647 (1:200; Thermo Scientific) for 30 minutes at room temperature. Finally, after

washing, samples were embedded in mowiol (Sigma Aldrich) and allowed to dry at RT. Confocal microscopy analysis was performed in a Leica SP8 microscope with a 63X oil objective (Leica Microsystems, Germany) using 405 nm (for CMAC), 488 nm (for Alexa-488) and 647 nm (for Alexa-647) excitation lines. Confocal sections were acquired every 0.25 µm along the z axis and 3D reconstructions were obtained with Image J software (National Institutes of Health). Graphs and statistics were made using PRISM 6 (GraphPad Software, USA).

### Cytotoxicity assay

CT26<sup>Luc</sup>, CT26<sup>EGFR-Luc</sup>, SW620<sup>Luc</sup> or HCT116<sup>Luc</sup> cells were plated in triplicates in 96-well microtiter plates at  $2 \times 10^4$  cells/well 24 hours before the assay. Then cells were incubated with CM or increasing amounts of purified antibodies for 30 min at 37°C. After that, Jurkat cells or human PBMCs were added at 5:1 E:T ratio. After 72 hours, specific cytotoxicity was determined adding the D-luciferin substrate (20 µg/mL, Promega) and relative light units (RLU) were measured with the luminescence plate reader Infinite 1200 (Tecan, Männedorf, Switzerland). Wells with target and effector cells in the absence of CM or purified antibodies were set as 100%. CM were also collected after 72 h hours and assayed for IFN-γ secretion by ELISA (cat#851.560, Diaclone).

### Inhibition of EGFR-mediated cell proliferation

A431 cells were seeded in triplicates at 2.000 cells/well in 96-well plates in DMEM supplemented with 10% FBS. After 24 h, medium was changed by DMEM containing 1% FBS and equimolar concentrations of purified AxOxE TriTE and OxE LiTE were added. Cetuximab and OKT3 were used as positive and negative controls, respectively. After 72 h, medium was removed and cell proliferation was measured adding CellTiter Glo luminescent assay (Promega, Madison, USA). Bioluminescence was assessed using a Tecan Infinite F200 plate-reading luminometer.

### EGFR signaling inhibition assay

A431 cells were seeded at 100.000 cells/well in 12-well plates in DMEM supplemented with 10% FBS and incubated for 24 hours. Afterward, cells were starved for 16 h with 1% FBS DMEM. Subsequently, cells were incubated with serum-free DMEM containing serial dilutions of AxOxE TriTE (200–0 nM) or OxE LiTE at 200 nM. Cetuximab was used as positive control. Then, cells were stimulated for 5 min with 25 ng/mL of human EGF and lysed in Laemmli-lysis buffer (Bio-Rad, CA, USA) for 10 min, on ice. Samples were analyzed by SDS-PAGE and Western blot using iBlot Dry Blotting System (Invitrogen Life Technologies). Membranes were incubated ON with a rabbit anti-human phospho EGFR Tyr1068 mAb (clone D7A5, Cell Signaling, Leiden, The Netherlands) and a anti β-actin mouse mAb (clone 8226, Abcam, Cambridge, UK), followed by incubation with an IRDye800-conjugated donkey anti-rabbit antibody (Rockland Immunochemicals, Limerick, PA, USA) and IRDye680-conjugated donkey anti-mouse

antibody (Rockland Immunochemicals). Odyssey infrared imaging system (LI-COR Biosciences, Lincoln, NE, USA) was used to visualize and analyzed protein bands.

### In vivo antitumoral effect

HCT116 ( $2 \times 10^6$  cells/mouse) in PBS mixed with 30% matrigel (BD Biosciences) were implanted s.c into the right dorsal space of 5-week-old female Hsd:ATHymic Nude-*Foxn1*<sup>nu</sup> mice. Tumors were measured three times a week with a calliper and their volumes estimated by using the formula: length x width<sup>2</sup> x 0.52, where length represents the largest tumor diameter and width represents the perpendicular tumor diameter. At day 4, mice were divided into groups with average diameter of 0.2 cm just before PBMCs administration. Randomization occurred in a blinded fashion. One day after intraperitoneal infusion of fresh PBMCs ( $1 \times 10^7$  cells/mouse) from a healthy donor (day 5), mice received i.p. injections of PBS, 60  $\mu$ g/mouse of OxE LiTE or 80  $\mu$ g/mouse of AxOxE TriTE (equimolar conditions), then treatment continued daily for another 9 days (until day 14). Mice were euthanized when tumor size reached 1 cm<sup>3</sup> or at the onset of any sign of distress. All experiments were conducted in compliance with the institutional guidelines provided by the Biomedical Research Institute Hospital Puerta de Hierro Animal Ethics Committee. Procedures were additionally approved by the Animal Welfare Division of the Environmental Affairs Council, Comunidad Autónoma de Madrid (PROEX 066/14).

### Histological studies

Mouse CRC xenografts from all mice were routinely formalin-fixed and paraffin-embedded in the Department of Pathology, Hospital Universitario Puerta de Hierro. Sections of 4  $\mu$ m were stained with hematoxylin and eosin according to standard protocols or processed for immunohistochemistry using the Dako-Omnis automated staining platform. The polyclonal rabbit anti-human CD3 ready-to-use (cat#GA503, Dako-Agilent) was developed using EnVision Flex High pH visualization system. At least two sections (three fields/section) of each tumor were blindly scored by the pathologist.

### Statistical analysis

Results were shown as mean  $\pm$  standard deviation (SD). Data were analyzed by unpaired two-tailed Student's t-test, assuming a normal distribution, using Prism software v5 (GraphPad, San Diego, CA, USA). Data were considered statistically significant when  $P < .05$ .

### Abbreviations:

BsAb, bispecific antibody; BiTE, bispecific T-cell engager; EGFR, epidermal growth factor receptor; EpCAM, epithelial cell adhesion molecule, PBMCs, peripheral blood mononuclear cells; scFv, single-chain variable fragment; TAA, tumor-associated antigen; TriKE, trispecific killer engager; TriTE, trispecific T-cell engager; TsAb, trispecific antibody; V<sub>HH</sub>,

immunoglobulin variable heavy chain; V<sub>HH</sub>, variable domain of heavy-chain only antibodies; V<sub>L</sub>, immunoglobulin variable light chain.

### Acknowledgments

The authors wish to thank donors and Biobank of Hospital Universitario Puerta de Hierro Majadahonda for the human specimens used in this study.

### Consent to Participate

Informed consent was obtained from all participants included in the study.

### Disclosure statement

No potential conflicts of interest were disclosed.

### Ethics Approval

All procedures involving animals were in accordance with the ethical standards of the corresponding institutional and regional/national committees.

### Funding

This study was funded by grants from Instituto de Salud Carlos III PI16/00357, PI19/00132), partially supported by the European Regional Development Fund (ERDF), Comunidad Autónoma de Madrid (S2010-BMD-2312), and Ministerio de Economía y Competitividad (RTC-2016-5118-1) to L.S.; and from Ministerio de Ciencia e Innovación (SAF2017-89437-P and PID2020-117323RB-I00), partially supported by ERDF, the Spanish Association Against Cancer (AECC 19084) and the CRIS Cancer Foundation FCRIS-2018-0042, FCRIS-2021-0090 (FCRIS-2018-0042 and FCRIS-2021-0090) to L.A.-V. A.T.-G. was supported by a predoctoral fellowship from Comunidad Autónoma de Madrid (PEJD-2018-PRE/BMD-8314); Spanish Ministry of Science and Innovation [SAF2017-89437-P, PID2020-117323RB-I00].

### ORCID

Antonio Tapia-Galisteo  <http://orcid.org/0000-0002-0507-8435>  
 Íñigo Sánchez Rodríguez  <http://orcid.org/0000-0002-6440-0922>  
 Oscar Aguilar-Sopeña  <http://orcid.org/0000-0002-2435-8598>  
 Seandean Lykke Harwood  <http://orcid.org/0000-0003-4654-8832>  
 Mariola Ferreras Gutierrez  <http://orcid.org/0000-0003-4421-3158>  
 Rocío Navarro  <http://orcid.org/0000-0002-0083-7711>  
 Cesáreo Corbacho  <http://orcid.org/0000-0002-6644-3475>  
 Marta Compte  <http://orcid.org/0000-0002-7138-9266>  
 Javier Lacadena  <http://orcid.org/0000-0002-7314-0333>  
 Francisco J. Blanco  <http://orcid.org/0000-0003-2545-4319>  
 Patrick Chames  <http://orcid.org/0000-0002-6104-6286>  
 Pedro Roda-Navarro  <http://orcid.org/0000-0003-3799-8823>  
 Luis Álvarez-Vallina  <http://orcid.org/0000-0003-3053-6757>  
 Laura Sanz  <http://orcid.org/0000-0002-3119-3218>

### References

- Kontermann RE, Brinkmann U. Bispecific Antibodies. *Drug Discov Today*. 2015;20(7):838–847. doi:10.1016/j.drudis.2015.02.008.
- Przepiorka D, Ko C-W, Deisseroth A, Yancey CL, Candau-Chacon R, Chiu H-J, Gehrke BJ, Gomez-Broughton C, Kane RC, Kirshner S, et al. FDA Approval: blinatumomab. *Clin Cancer Res*. 2015;21(18):4035–4039. doi:10.1158/1078-0432.CCR-15-0612.

3. Suurs FV, Lub-de Hooge MN, de Vries EGE, de Groot DJA. A Review of Bispecific Antibodies and Antibody Constructs in Oncology and Clinical Challenges. *Pharmacol Ther.* 2019;201:103–119. doi:10.1016/j.pharmthera.2019.04.006.
4. Compte M, Harwood SL, Muñoz IG, Navarro R, Zonca M, Perez-Chacon G, Erce-Llamazares A, Merino N, Tapia-Galisteo A, Cuesta AM, et al. A Tumor-Targeted Trimeric 4-1BB-Agonistic Antibody Induces Potent Anti-Tumor Immunity without Systemic Toxicity. *Nat Commun.* 2018;9(1):4809. doi:10.1038/s41467-018-07195-w.
5. Gauthier L, Morel A, Anceriz N, Rossi B, Blanchard-Alvarez A, Grondin G, Trichard S, Cesari C, Sapet M, Bosco F, et al. Multifunctional Natural Killer Cell Engagers Targeting Nkp46 Trigger Protective Tumor Immunity. *Cell.* 2019;177(7):1701–1713.e16. doi:10.1016/j.cell.2019.04.041.
6. Wu L, Seung E, Xu L, Rao E, Lord DM, Wei RR, Cortez-Retamozo V, Ospina B, Posternak V, Uliniski G, et al. Trispecific Antibodies Enhance the Therapeutic Efficacy of Tumor-Directed T Cells through T Cell Receptor Co-Stimulation. *Nat Cancer.* 2020;1(1):86–98. doi:10.1038/s43018-019-0004-z.
7. Moore GL, Bennett MJ, Rashid R, Pong EW, Nguyen D-HT, Jacinto J, Eivazi A, Nisthal A, Diaz JE, Chu SY, et al. A Robust Heterodimeric Fc Platform Engineered for Efficient Development of Bispecific Antibodies of Multiple Formats. *Methods.* 2019;154:38–50. doi:10.1016/j.ymeth.2018.10.006.
8. Herrmann M, Krupka C, Deiser K, Brauchle B, Marcinek A, Ogrinc Wagner A, Rataj F, Mocikat R, Metzler KH, Spiekermann K, et al. Bifunctional PD-1 × ACD3 × ACD33 Fusion Protein Reverses Adaptive Immune Escape in Acute Myeloid Leukemia. *Blood.* 2018;132(23):2484–2494. doi:10.1182/blood-2018-05-849802.
9. Braciak TA, Roskopf CC, Wildenhain S, Fenn NC, Schiller CB, Schubert IA, Jacob U, Honegger A, Krupka C, Subklewe M, et al. Dual-Targeting Triplebody 33-16-123 (SPM-2) Mediates Effective Redirected Lysis of Primary Blasts from Patients with a Broad Range of AML Subtypes in Combination with Natural Killer Cells. *Oncoimmunology.* 2018;7(9):e1472195. doi:10.1080/2162402X.2018.1472195.
10. Schubert I, Kellner C, Stein C, Kugler M, Schwenkert M, Saul D, Mentz K, Singer H, Stockmeyer B, Hillen W, et al. A Single-Chain Triplebody with Specificity for CD19 and CD33 Mediates Effective Lysis of Mixed Lineage Leukemia Cells by Dual Targeting. *MABS.* 2011;3(1):21–30. doi:10.4161/mabs.3.1.14057.
11. Kugler M, Stein C, Kellner C, Mentz K, Saul D, Schwenkert M, Schubert I, Singer H, Oduncu F, Stockmeyer B, et al. A Recombinant Trispecific Single-Chain Fv Derivative Directed against CD123 and CD33 Mediates Effective Elimination of Acute Myeloid Leukemia Cells by Dual Targeting. *Br J Haematol.* 2010;150(5):574–586. doi:10.1111/j.1365-2141.2010.08300.x.
12. Roskopf CC, Schiller CB, Braciak TA, Kobold S, Schubert IA, Fey GH, Hopfner K-P, Oduncu FS. T Cell-Recruiting Triplebody 19-3-19 Mediates Serial Lysis of Malignant B-Lymphoid Cells by a Single T Cell. *Oncotarget.* 2014;5(15):6466–6483. doi:10.18632/oncotarget.2238.
13. Roskopf CC, Braciak TA, Fenn NC, Kobold S, Fey GH, Hopfner K-P, Oduncu FS. Dual-Targeting Triplebody 33-3-19 Mediates Selective Lysis of Biphenotypic CD19+ CD33+ Leukemia Cells. *Oncotarget.* 2016;7(16):22579–22589. doi:10.18632/oncotarget.8022.
14. Felices M, Kodal B, Hinderlie P, Kaminski MF, Cooley S, Weisdorf DJ, Vallera DA, Miller JS, Bachanova V. Novel CD19-Targeted TriKE Restores NK Cell Function and Proliferative Capacity in CLL. *Blood Adv.* 2019;3(6):897–907. doi:10.1182/bloodadvances.2018029371.
15. Vallera DA, Felices M, McElmurry R, McCullar V, Zhou X, Schmohl JU, Zhang B, Lenvik AJ, Panoskaltis-Mortari A, Vermeris MR, et al. IL15 Trispecific Killer Engagers (TriKE) Make Natural Killer Cells Specific to CD33+ Targets While Also Inducing Persistence, In Vivo Expansion, and Enhanced Function. *Clin Cancer Res.* 2016;22(14):3440–3450. doi:10.1158/1078-0432.CCR-15-2710.
16. Arvindam US, van Hauten PMM, Schirm D, Schaap N, Hobo W, Blazar BR, Vallera DA, Dolstra H, Felices M, Miller JS. A Trispecific Killer Engager Molecule against CLEC12A Effectively Induces NK-Cell Mediated Killing of AML Cells. *Leukemia.* 2021;35(6):1586–1596. doi:10.1038/s41375-020-01065-5.
17. Austin RJ, Lemon BD, Aaron WH, Barath M, Culp PA, DuBridge RB, Evnin LB, Jones A, Panchal A, Patnaik P, et al. TriTACs, a Novel Class of T-Cell-Engaging Protein Constructs Designed for the Treatment of Solid Tumors. *Mol Cancer Ther.* 2021;20(1):109–120. doi:10.1158/1535-7163.MCT-20-0061.
18. Catalano I, Trusolino L. The Stromal and Immune Landscape of Colorectal Cancer Progression during Anti-EGFR Therapy. *Cancer Cell.* 2019;36(1):1–3. doi:10.1016/j.ccell.2019.06.001.
19. Gires O, Pan M, Schinke H, Canis M, Baeuerle PA. Expression and Function of Epithelial Cell Adhesion Molecule EpCAM: where Are We after 40 Years? *Cancer Metastasis Rev.* 2020;39(3):969–987. doi:10.1007/s10555-020-09898-3.
20. Borlak J, Länger F, Spanel R, Schöndorfer G, Dittrich C. Immune-Mediated Liver Injury of the Cancer Therapeutic Antibody Catumaxomab Targeting EpCAM, CD3 and Fcγ Receptors. *Oncotarget.* 2016;7(19):28059–28074. doi:10.18632/oncotarget.8574.
21. Even-Desrumeaux K, Nevoltris D, Lavaut MN, Alim K, Borg J-P, Audebert S, Kerfelec B, Baty D, Chames P. Masked Selection: a Straightforward and Flexible Approach for the Selection of Binders against Specific Epitopes and Differentially Expressed Proteins by Phage Display. *Mol Cell Proteomics.* 2014;13(2):653–665. doi:10.1074/mcp.O112.025486.
22. Hofman EG, Ruonala MO, Bader AN, van den Heuvel D, Voortman J, Roovers RC, Verkleij AJ, Gerritsen HC, van Bergen En Henegouwen PMP. EGF Induces Coalescence of Different Lipid Rafts. *J Cell Sci.* 2008;121(Pt 15):2519–2528. doi:10.1242/jcs.028753.
23. Holliger P, Manzke O, Span M, Hawkins R, Fleischmann B, Qinghua L, Wolf J, Diehl V, Cochet O, Winter G, et al. Carcinoembryonic Antigen (CEA)-Specific T-Cell Activation in Colon Carcinoma Induced by Anti-CD3 × Anti-CEA Bispecific Diabodies and B7 × Anti-CEA Bispecific Fusion Proteins. *Cancer Res.* 1999;59:2909–2916.
24. Mølgaard K, Harwood SL, Compte M, Merino N, Bonet J, Alvarez-Cienfuegos A, Mikkelsen K, Nuñez-Prado N, Alvarez-Mendez A, Sanz L, et al. Bispecific Light T-Cell Engagers for Gene-Based Immunotherapy of Epidermal Growth Factor Receptor (EGFR)-Positive Malignancies. *Cancer Immunol Immunother.* 2018;67(8):1251–1260. doi:10.1007/s00262-018-2181-5.
25. Pozzi C, Cuomo A, Spadoni I, Magni E, Silvola A, Conte A, Sigismund S, Ravenda PS, Bonaldi T, Zampino MG, et al. The EGFR-Specific Antibody Cetuximab Combined with Chemotherapy Triggers Immunogenic Cell Death. *Nat Med.* 2016;22(6):624–631. doi:10.1038/nm.4078.
26. Frenzel A, Hust M, Schirrmann T. Expression of Recombinant Antibodies. *Front Immunol.* 2013;4:217. doi:10.3389/fimmu.2013.00217.
27. Zheng S, Moores S, Jarantow S, Pardinias J, Chiu M, Zhou H, Wang W. Cross-Arm Binding Efficiency of an EGFR x c-Met Bispecific Antibody. *MABS.* 2016;8(3):551–561. doi:10.1080/19420862.2015.1136762.
28. Harms BD, Kearns JD, Iadevaia S, Lugovskoy AA. Understanding the Role of Cross-Arm Binding Efficiency in the Activity of Monoclonal and Multispecific Therapeutic Antibodies. *Methods.* 2014;65(1):95–104. doi:10.1016/j.ymeth.2013.07.017.
29. Schmitz KR, Bagchi A, Roovers RC, van Bergen En Henegouwen PMP, Ferguson KM. Structural Evaluation of EGFR Inhibition Mechanisms for Nanobodies/VHH Domains. *Structure.* 2013;21(7):1214–1224. doi:10.1016/j.str.2013.05.008.

30. Hristodorov D, Amoury M, Mladenov R, Niesen J, Arens K, Berges N, Hein L, Di Fiore S, Pham A-T, Huhn M, et al. EpCAM-Selective Elimination of Carcinoma Cells by a Novel MAP-Based Cytolytic Fusion Protein. *Mol Cancer Ther.* 2014;13(9):2194–2202. doi:10.1158/1535-7163.MCT-13-0781.
31. Harwood SL, Alvarez-Cienfuegos A, Nuñez-Prado N, Compte M, Hernández-Pérez S, Merino N, Bonet J, Navarro R, Van Bergen En Henegouwen PMP, Lykkemark S, et al. ATTACK, a Novel Bispecific T Cell-Recruiting Antibody with Trivalent EGFR Binding and Monovalent CD3 Binding for Cancer Immunotherapy. *Oncoimmunology.* 2017;7(1):e1377874. doi:10.1080/2162402X.2017.1377874.
32. Marusyk A, Janiszewska M, Polyak K. Intratumor Heterogeneity: the Rosetta Stone of Therapy Resistance. *Cancer Cell.* 2020;37(4):471–484. doi:10.1016/j.ccell.2020.03.007.
33. Gonzalez-Exposito R, Semiannikova M, Griffiths B, Khan K, Barber LJ, Woolston A, Spain G, von Loga K, Challoner B, Patel R, et al. CEA Expression Heterogeneity and Plasticity Confer Resistance to the CEA-Targeting Bispecific Immunotherapy Antibody Cibisatamab (CEA-TCB) in Patient-Derived Colorectal Cancer Organoids. *J Immunother Cancer.* 2019;7(1):101. doi:10.1186/s40425-019-0575-3.
34. Ruella M, Barrett DM, Kenderian SS, Shestova O, Hofmann TJ, Perazzelli J, Klichinsky M, Aikawa V, Nazimuddin F, Kozlowski M, et al. Dual CD19 and CD123 Targeting Prevents Antigen-Loss Relapses after CD19-Directed Immunotherapies. *J Clin Invest.* 2016;126(10):3814–3826. doi:10.1172/JCI87366.
35. Montagut C, Dalmases A, Bellosillo B, Crespo M, Pairet S, Iglesias M, Salido M, Gallen M, Marsters S, Tsai SP, et al. Identification of a Mutation in the Extracellular Domain of the Epidermal Growth Factor Receptor Confering Cetuximab Resistance in Colorectal Cancer. *Nat Med.* 2012;18(2):221–223. doi:10.1038/nm.2609.
36. Bertotti A, Papp E, Jones S, Adleff V, Anagnostou V, Lupo B, Sausen M, Phallen J, Hruban CA, Tokheim C, et al. The Genomic Landscape of Response to EGFR Blockade in Colorectal Cancer. *Nature.* 2015;526(7572):263–267. doi:10.1038/nature14969.
37. Metges J, Ramée JF, Dupuis O, Deguiral P, Boucher E, Cojocarasu O, Ferec M, Porneuf M, Douillard J, Grude F. Panerb Study: which Category of Patients, Suffering from Metastatic Colorectal Cancer, Can Benefit From Panitumumab Treatment After Cetuximab-Based Regimen Failure? *Annals of Oncology.* 2012;23:ix196. doi:10.1016/S0923-7534(20)33183-5.
38. Arena S, Bellosillo B, Siravegna G, Martínez A, Cañadas I, Lazzari L, Ferruz N, Russo M, Misale S, González I, et al. Emergence of Multiple EGFR Extracellular Mutations during Cetuximab Treatment in Colorectal Cancer. *Clin Cancer Res.* 2015;21(9):2157–2166. doi:10.1158/1078-0432.CCR-14-2821.
39. Goebeler M-E, Bargou RC. T Cell-Engaging Therapies - BiTEs and Beyond. *Nat Rev Clin Oncol.* 2020;17(7):418–434. doi:10.1038/s41571-020-0347-5.
40. Shah NN, Johnson BD, Schneider D, Zhu F, Szabo A, Keever-Taylor CA, Krueger W, Worden AA, Kadan MJ, Yim S, et al. Bispecific Anti-CD20, Anti-CD19 CAR T Cells for Relapsed B Cell Malignancies: a Phase 1 Dose Escalation and Expansion Trial. *Nat Med.* 2020;26(10):1569–1575. doi:10.1038/s41591-020-1081-3.
41. Rossotti MA, Bélanger K, Henry KA, Tanha J. Immunogenicity and Humanization of Single-Domain Antibodies. *FEBS J.* 2021. doi:10.1111/febs.15809.
42. Bannas P, Hambach J, Koch-Nolte F. Nanobodies and Nanobody-Based Human Heavy Chain Antibodies As Antitumor Therapeutics. *Front Immunol.* 2017;8:1603. doi:10.3389/fimmu.2017.01603.
43. Punt CJ, Nagy A, Douillard J-Y, Figer A, Skovsgaard T, Monson J, Barone C, Fountzilias G, Riess H, Moylan E, et al. Edrecolomab Alone or in Combination with Fluorouracil and Folinic Acid in the Adjuvant Treatment of Stage III Colon Cancer: a Randomised Study. *The Lancet.* 2002;360(9334):671–677. doi:10.1016/S0140-6736(02)09836-7.
44. Goel S, Bauer RJ, Desai K, Bulgaru A, Iqbal T, Strachan B-K, Kim G, Kaubisch A, Vanhove GF, Goldberg G, et al. Pharmacokinetic and Safety Study of Subcutaneously Administered Weekly ING-1, a Human Engineered Monoclonal Antibody Targeting Human EpCAM, in Patients with Advanced Solid Tumors. *Ann Oncol.* 2007;18(10):1704–1707. doi:10.1093/annonc/mdm280.
45. Schmidt M, Scheulen ME, Dittrich C, Obrist P, Marschner N, Dirix L, Schmidt M, Rüttinger D, Schuler M, Reinhardt C, et al. An Open-Label, Randomized Phase II Study of Adecatumumab, a Fully Human Anti-EpCAM Antibody, as Monotherapy in Patients with Metastatic Breast Cancer. *Ann Oncol.* 2010;21(2):275–282. doi:10.1093/annonc/mdp314.
46. Jarantow SW, Bushey BS, Pardinas JR, Boakye K, Lacy ER, Sanders R, Sepulveda MA, Moores SL, Chiu ML. Impact of Cell-Surface Antigen Expression on Target Engagement and Function of an Epidermal Growth Factor Receptor × c-MET Bispecific Antibody. *J Biol Chem.* 2015;290(41):24689–24704. doi:10.1074/jbc.M115.651653.
47. Zuckier LS, Berkowitz EZ, Sattenberg RJ, Zhao QH, Deng HF, Scharff MD. Influence of Affinity and Antigen Density on Antibody Localization in a Modifiable Tumor Targeting Model. *Cancer Res.* 2000;60:7008–7013.
48. Mazor Y, Oganeyan V, Yang C, Hansen A, Wang J, Liu H, Sachsenmeier K, Carlson M, Gadre DV, Borrok MJ, et al. Improving Target Cell Specificity Using a Novel Monovalent Bispecific IgG Design. *MAbs.* 2015;7(2):377–389. doi:10.1080/19420862.2015.1007816.
49. Mazor Y, Sachsenmeier KF, Yang C, Hansen A, Filderman J, Mulgrew K, Wu H, Dall'Acqua WF. Enhanced Tumor-Targeting Selectivity by Modulating Bispecific Antibody Binding Affinity and Format Valence. *Sci Rep.* 2017;7(1):40098. doi:10.1038/srep40098.
50. Navarro R, Tapia-Galisteo A, Martín-García L, Tarín C, Corbacho C, Gómez-López G, Sánchez-Tirado E, Campuzano S, González-Cortés A, Yáñez-Sedeño P, et al. TGF-β-Induced IGFBP-3 Is a Key Paracrine Factor from Activated Pericytes That Promotes Colorectal Cancer Cell Migration and Invasion. *Mol Oncol.* 2020;14(10):2609–2628. doi:10.1002/1878-0261.12779.

## A PD-L1/EGFR bispecific antibody combines immune checkpoint blockade and direct anti-cancer action for an enhanced anti-tumor response

Laura Rubio-Pérez <sup>a,b,c,d</sup>, Rodrigo Lázaro-Gorines <sup>a,b,c</sup>, Seandean L. Harwood <sup>e</sup>, Marta Compte <sup>f</sup>, Rocío Navarro <sup>f</sup>, Antonio Tapia-Galisteo <sup>a,b,c</sup>, Jaime Bonet <sup>g</sup>, Belén Blanco <sup>a,b,c</sup>, Simon Lykkemark <sup>h</sup>, Ángel Ramírez-Fernández <sup>a,b,c</sup>, Mariola Ferreras-Gutiérrez <sup>i</sup>, Carmen Domínguez-Alonso <sup>a,b,c</sup>, Laura Díez-Alonso <sup>a,b,c</sup>, Alejandro Segura-Tudela <sup>a,b,c</sup>, Oana Hangiu <sup>a,b,f</sup>, Ainhoa Erce-Llamazares <sup>a,b,f</sup>, Francisco J. Blanco <sup>d</sup>, Cruz Santos <sup>d</sup>, José L. Rodríguez-Peralto <sup>k,l,m,n</sup>, Laura Sanz <sup>n,o</sup>, and Luis Álvarez-Vallina <sup>a,b,c,d</sup>

<sup>a</sup>Cancer Immunotherapy Unit (UNICA), Department of Immunology, Hospital Universitario 12 de Octubre, Madrid, Spain; <sup>b</sup>Immuno-Oncology and Immunotherapy Group, Instituto de Investigación Sanitaria 12 de Octubre (imas12), Madrid, Spain; <sup>c</sup>H120-CNIO Cancer Immunotherapy Clinical Research Unit, Spanish National Cancer Research Centre (CNIO), Madrid, Spain; <sup>d</sup>Chair for Immunology UFV/Merck, Universidad Francisco de Vitoria (UFV), Madrid, Spain; <sup>e</sup>Department of Molecular Biology and Genetics, Aarhus University, Aarhus C, Denmark; <sup>f</sup>Department of Antibody Engineering, Leadartis SL, QUBE Technology Park, Madrid, Spain; <sup>g</sup>Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland; <sup>h</sup>Immunotherapy and Cell Engineering Laboratory, Department of Engineering, Aarhus University, Aarhus C, Denmark; <sup>i</sup>Centro de Investigaciones Biológicas Margarita Salas (CIB), CSIC, Madrid, Spain; <sup>j</sup>Faculty of Experimental Sciences, Universidad Francisco de Vitoria (UFV), Madrid, Spain; <sup>k</sup>Department of Pathology, Hospital Universitario 12 de Octubre, Madrid, Spain; <sup>l</sup>Department of Pathology, Universidad Complutense, Madrid, Spain; <sup>m</sup>Cutaneous Oncology Group, Instituto de Investigación Sanitaria 12 de Octubre (imas12), Madrid, Spain; <sup>n</sup>Centro de Investigación Biomédica en Red en Oncología (CIBERONC), Madrid, Spain; <sup>o</sup>Molecular Immunology Unit, Hospital Universitario Puerta de Hierro Majadahonda, Madrid, Spain

### ABSTRACT

Immune checkpoint blockade (ICB) with antibodies has shown durable clinical responses in a wide range of cancer types, but the overall response rate is still limited. Other effective therapeutic modalities to increase the ICB response rates are urgently needed. New bispecific antibody (bsAb) formats combining the ICB effect and a direct action on cancer cells could improve the efficacy of current immunotherapies. Here, we report the development of a PD-L1/EGFR symmetric bsAb by fusing a dual-targeting tandem trimmer body with the human IgG1 hinge and Fc regions. The bsAb was characterized *in vitro* and the antitumor efficacy was evaluated in humanized mice bearing xenografts of aggressive triple-negative breast cancer and lung cancer. The IgG-like hexavalent bsAb, designated IgTT-1E, was able to simultaneously bind both EGFR and PD-L1 antigens, inhibit EGF-mediated proliferation, effectively block PD-1/PD-L1 interaction, and induce strong antigen-specific antibody-dependent cellular cytotoxicity activity *in vitro*. Potent therapeutic efficacies of IgTT-1E in two different humanized mouse models were observed, where tumor growth control was associated with a significantly increased proportion of CD8<sup>+</sup> T cells. These results support the development of IgTT-1E for the treatment of EGFR<sup>+</sup> cancers.

### ARTICLE HISTORY

Received 2 January 2023  
Revised 16 April 2023  
Accepted 17 April 2023

### KEYWORDS



Cancer immunotherapy, bispecific antibody; dual action; epithelial growth factor receptor; immune checkpoint blockade


## Introduction

Modulating immune responses using monoclonal antibodies (mAbs) is one of the most promising approaches for cancer immunotherapy<sup>1</sup>. Immune checkpoints are coinhibitory and costimulatory receptors that are crucial for maintaining self-tolerance and modulating immune responses<sup>2</sup>. However, inhibitory checkpoints can be “hijacked” by tumors to evade immune responses<sup>3</sup>. mAb-based blockade of coinhibitory immune receptors, such as cytotoxic T lymphocyte antigen 4 (CTLA-4), programmed cell death-1 (PD-1) receptor, or PD-1 ligand (PD-L1) has shown durable tumor clinical responses in a wide range of cancer types, but their efficacy is limited to 10% to 30% of patients<sup>4</sup>. Different factors, such as the tumor mutational burden and immunogenicity of cancer cells, the composition of the tumor microenvironment (TME) and the degree of immune cell infiltration, influence the response to immune

checkpoint blockade (ICB)<sup>5</sup>. Increasing the response rates to ICB is likely to require the design of therapeutic combinations that are tailored to the aforementioned factors<sup>6</sup>. On the other hand, EGFR is a receptor tyrosine kinase well-known as a tumor-associated antigen (TAA) with an important functional role, since EGFR deregulation promotes proliferation, apoptosis inhibition and invasion of cancer cells<sup>7</sup>. Therefore, mAb targeting EGFR not only mark tumor cells for immune killing but also may inhibit phosphorylation of key intracellular tyrosine residues, thus preventing ligand-mediated mitogenic signaling<sup>8</sup>.

As of January 2023, seven immune checkpoint blockers (ICB) preventing PD1/PD-L1 interaction had been approved for clinical use by the Food and Drug Administration (FDA) and the European Medicines Agency (EMA), and six others are in regulatory review, being *atezolizumab* the first anti-PD-L1

**CONTACT** Luis Álvarez-Vallina  [lav.imas12@h12o.es](mailto:lav.imas12@h12o.es)  Cancer Immunotherapy Unit (UNICA), Department of Immunology, Hospital Universitario 12 de Octubre, Avda. Córdoba s/n, Madrid 28041, Spain

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/2162402X.2023.2205336>

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

mAb in the market (2017). With respect to EGFR, three conventional mAbs [*cetuximab*, *panitumumab* –both approved for colorectal cancer (CRC)- and *nectinmumab* -for non-small cell lung cancer (NSCLC)- and the bispecific antibody (bsAb) *amivantamab* (EGFR x c-MET, for NSCLC with EGFR exon 20 insertion mutations) are currently in the market. Antibody therapeutics that are approved for marketing in regions other than the US or EU include two anti-EGFR (*nimotuzumab* and *cetuximab* conjugated to the photosensitizer IR700) and seven anti-PD1<sup>9</sup>.

Designing agents that combine the immunomodulatory effect of an ICB and a direct action on cancer cells could open new perspectives in cancer immunotherapy. Here, we generated a novel PD-L1/EGFR bispecific antibody (bsAb) by fusing a dual-targeting tandem trimerbody (TT)<sup>10</sup> with the human IgG1 hinge and Fc regions. The TT format has been previously described and is a single-chain fusion of three single-domain V<sub>HH</sub> antibodies with three collagen XVIII trimerization domains (TIE)<sup>8-10</sup>, with intercalating glycine-serine-based linkers. Each V<sub>HH</sub> is encoded separately, allowing mono-, bi-, and tri-specific molecules to be produced<sup>7</sup>. Here, the first V<sub>HH</sub> was replaced by a PD-L1-specific scFv, and an EGFR-specific V<sub>HH</sub> was used in the second and third V<sub>HH</sub> positions. The resulting IgG-like hexavalent bsAb, designated IgTT-1E, simultaneously bound both targeted antigens, inhibited EGF-mediated proliferation, effectively blocked PD-1/PD-L1 interaction and induced potent antigen-specific antibody-dependent cellular cytotoxicity (ADCC) activity *in vitro*. A potent therapeutic effect of IgTT-1E was observed in humanized mice bearing aggressive EGFR<sup>+</sup>PD-L1<sup>+</sup> human triple-negative breast cancer (TNBC) and lung cancer cell-line –derived xenografts (CDX). These data provide a promising basis for the further clinical development of IgTT-1E and similarly designed symmetric antibodies against different targets.

## Materials and methods

### Cell lines and culture conditions

HEK-293 (CRL-1573), A431 (CRL-1555), NIH/3T3 (CRL-1658), MDA-MB-231 (HTB-26) and CHO-K1 (CCL-61) cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Life Technologies, cat# 10313021) supplemented with 2 mmol/l L-glutamine, 10% (v/v) heat-inactivated fetal bovine serum (FBS) (Merck Life Science, cat# F7524-500 ML), and antibiotics (100 units/mL penicillin, 100 mg/mL streptomycin; both from Life Technologies) at 37 °C in 5% CO<sub>2</sub> humidity. A549 (CCL-185) cells were cultured in RPMI-1640 (Lonza, cat# 12-702Q) supplemented with 2 mmol/l L-glutamine, 10% (v/v) heat-inactivated FBS, and antibiotics. All these cell lines were obtained from the American Type Culture Collection. NIH/3T3 cells expressing human EGFR (3T3<sup>EGFR</sup>) were kindly provided by Dr. A. Villalobo [Instituto de Investigaciones Biomédicas “Alberto Sols”. IIBm (CSIC-UAM), Madrid, Spain]. Jurkat T cells stably expressing human PD-1 and NFAT-induced luciferase (Jurkat<sup>NFAT-RE-luc/PD-1</sup>) and CHO-K1 cells stably expressing human PD-L1 (PD-L1 aAPC/CHO-K1), were obtained from Promega (cat# J1250). CHO-K1 cells stably expressing

human PD-L1 (CHO<sup>PD-L1</sup>) were obtained from Genlantis (xCELLerate<sup>TM</sup> PD-L1 Stable Cell Line, XCL-PDL1) and CHO-K1 cells stably expressing human EGFR (CHO<sup>EGFR</sup>), or both (CHO<sup>EGFR-PD-L1</sup>) were generated using human EGFR encoding commercial lentiviral particles (G&P Biosciences, cat# LTV0169). All cell lines were routinely screened for mycoplasma contamination by PCR using the Mycoplasma Plus TM Primer Set (Biotools B&M Labs, cat# 90022).

### Construction of expression vectors

The plasmid pCR3.1-FLAG/Strep-αEGFR<sub>3</sub>-TIE-αCD3 was generated by cloning the insert OncoM-FLAG/Strep-αEGFR flanked by *HindIII-NotI* (GeneArt AG, Thermo Fischer) into the plasmid pCR3.1-αEGFR<sub>3</sub>-TIE-αCD3 encoding the bispecific EGFR x anti-CD3 ATTACK<sup>11</sup>. Then, the FLAG/Strep-αEGFR V<sub>HH</sub>-based tandem trimerbody (TT) was subcloned as *HindIII/BamHI* into pCR3.1-hFc-His vector containing a human IgG<sub>1</sub> hinge and Fc region, resulting in pCR3.1-FLAG/Strep-αEGFR<sub>3</sub>-hFc-His (IgTT-E). To generate the PD-L1/EGFR bispecific IgTT (IgTT-1E) expressing vector, the OncoM-FLAG/Strep-αPD-L1 fragment flanked by *HindIII-NotI* (GeneArt AG) was cloned into the plasmid pCR3.1-FLAG/Strep-αEGFR<sub>3</sub>-hFc-His. All the sequences were verified using primers FwCMV and RvBGH oligonucleotides (Table S1).

### Expression and purification of recombinant antibodies

HEK-293 cells were transfected with the appropriated vectors by Lipofectamine 3,000 transfection kit (Fisher Scientific, cat# 15292465) and selected in complete DMEM supplemented with 500 µg/ml of G418 to generate stable cell lines. Conditioned media were collected and processed using Strep-Tactin purification system (IBA Lifesciences) in an ÄKTA Prime plus system (Life Technologies). The purified antibodies were dialyzed overnight at 4 °C against PBS pH 7.4 supplemented with 150 mM NaCl and analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) under reducing conditions.

### Western blotting

Protein samples were analyzed by mean of 10–20% Tris-glycine SDS-PAGE under reducing conditions, transferred onto nitrocellulose membranes (Thermo Fisher Scientific, cat# IB23002) and probed with mouse anti-FLAG IgG<sub>1</sub> (clone M2, Sigma-Aldrich, cat# F3165) (1 µg/ml), followed by incubation with HRP-conjugated goat anti-mouse IgG (GAM-HRP) (1:10,000 dilution) (Sigma, cat# A2554). Visualization of protein bands was performed with Pierce ECL Plus Western Blotting Substrate (Thermo Scientific, cat# 32132), using ChemiDoc MP Imaging System and Image Lab software (both from BioRad).

### Enzyme-linked immunosorbent assay

The human EGFR-Fc (EGFR-Fc, R&D Systems, cat# 344-ER) or human PD-L1-Fc (PD-L1-Fc, Peprotech, cat# 310-35)



chimeras were immobilized (2.5 µg/ml in PBS) on Maxisorp 96-well plates (NUNC Brand Products, cat# 44240) overnight at 4°C. After washing and blocking, conditioned media or purified protein solution (1 µg/ml) was added and incubated for 1 hour at room temperature. The wells were washed and HRP-conjugated anti-poly Histidine (Sigma-Aldrich, cat# A7058), HRP-conjugated anti-FLAG (M2 clone, Sigma-Aldrich, cat# A8592), mouse anti-Myc (clone 9E10, Millipore, cat# 05-419) or HRP-conjugated goat anti-human IgG (GAH) (Sigma-Aldrich, cat# A0170) were added (1 µg/ml). After washing, in the case of mouse anti-Myc, GAM-HRP (1:2,000 dilution) (Jackson ImmunoResearch, cat# 115-085-166) was added for 1 hour at room temperature. Finally, after washing, the plate was developed using 100 µl 3,3',5,5'-tetramethylbenzidine (TMB) (Sigma-Aldrich, cat# T0440) and stopped by 100 µl of 1 N H<sub>2</sub>SO<sub>4</sub>. Absorbance was read at 450–620 nm using Multiskan FC photometer (Thermo Scientific).

### Size exclusion chromatography-multiangle light scattering

The experiments were performed on a Superdex 200 Increase 10/300 GL column (Cytiva) attached in-line to a DAWN EOS light scattering photometer (Wyatt Technology) and an ultraviolet light absorbance detector (ThermoFinnigan SpectraSYSTEM UV2000). The chromatography was run at room temperature and the scattering detector was thermostated at 23 °C. The column was equilibrated with running buffer (PBS pH 7.4 plus 150 mM NaCl, 0.1 µm filtered) and the size exclusion chromatography-multiangle light scattering (SEC-MALS) system was calibrated with a sample of BSA (Albumin, Monomer bovine, Sigma-Aldrich, cat#A1900) at 2 mg/ml in the same buffer. 230 µl of the antibody solutions at 0.3 or 0.4 mg/ml were injected into the column at a flow rate of 0.5 ml/min. The column has an exclusion volume of 8.6 ml, and no absorbance (no aggregated proteins) was observed in the chromatograms at this volume. Data acquisition and analysis were performed using ASTRA software (Wyatt Technology). The reported molar masses correspond to the center of the chromatography peaks. Based on numerous measurements on BSA samples under similar conditions the estimation of the experimental error in the molar mass is around 5%.

### Molecular Modeling

The IgTT-E three-dimensional representation was built by homology modeling using MODELLER<sup>12</sup>. Two templates were combined to generate the model, the V<sub>HH</sub>-based monospecific TT was built using as template the anti-CEA TT constructed in a previous work<sup>13</sup>, while the Fc domain was based on the human IgG1 B12 structure (pdb:1HZH.H)<sup>14</sup>, obtained from the Protein Data Bank<sup>15</sup>. The Fc template was obtained with BLAST<sup>16</sup>, with an e-value of 1e-169 and a 99% of sequence identity for the domain. The structure of 1HZH was also used to guide dimerization. The IgTT-1E model was built by combining the IgTT-E model and the interleukin 18 receptor antagonist scFv (pdb:6NK9.D)<sup>17</sup> for the anti-PD-L1

scFv domain. The template was obtained through BLAST with an e-value of 2e-138 and 83% of sequence identity for the domain.

### Biolayer interferometry

The binding of the IgTT-E to immobilized EGFR-Fc, and of the IgTT-1E to immobilized EGFR-Fc and PD-L1-Fc was measured using biolayer interferometry (BLI) on an Octet RED96 system (Fortebio). The binding to EGFR was compared to that of *cetuximab* (ctx) (Merck KGaA, Darmstadt, Germany), and the binding to PD-L1 was compared with *atezolizumab* (atz) (Fritz Hoffmann-La Roche, Basel, Switzerland). Both antigens were immobilized onto AR2G biosensors (Fortebio) at pH 5.0 using amine reactive coupling. Antibodies in HEPES-buffered saline (HBS; 20 mM HEPES, 150 mM NaCl, pH 7.4) at 10 nM were associated with either immobilized antigen for 30 minutes, after which the dissociation of antibody from the biosensor was measured for 30 minutes in HBS buffer only. After the IgTT-1E and ctx had bound to immobilized EGFR, bispecific binding to PD-L1-Fc in solution was then investigated by the subsequent treatment of the biosensors with 20 nM of PD-L1-Fc in HBS for 30 minutes. To determine binding kinetics to immobilized EGFR and PD-L1, the Octet Data Analysis (Fortebio) software was used to fit the experiment data to a 1:1 binding model with an asymptotic association phase and an exponential decay dissociation phase and regression to the chi<sup>2</sup> value.

### Serum stability

Purified IgTT-1E was incubated in 60% (v/v) human serum (Sigma-Aldrich, cat# H4522) at 37°C for 96 hours. The binding activity of the sample at 0 time point was set as 100% to calculate the corresponding decay in PD-L1 and EGFR binding by enzyme-linked immunosorbent assay (ELISA). Samples were analyzed with Multiskan FC Photometer and GraphPad Prism software. Results correspond to one experiment performed in triplicate.

### Flow cytometry

CHO<sup>EGFR</sup> or CHO<sup>PD-L1</sup> cells (1 × 10<sup>5</sup> cells/well) were incubated for 1 hour on ice with purified antibodies (6.67 nM), washed and incubated for 30 minutes with a PE-conjugated F(ab')<sub>2</sub> GAH IgG antibody (Jackson ImmunoResearch, cat# 109-116-170). *Trastuzumab* (tra) (Fritz Hoffmann-La Roche), atz and ctx (6.67 nM) were used as controls. After washing, DAPI (Sigma Aldrich, cat# D9542) was added and samples were analyzed with FACSCanto II Flow Cytometer (Becton Dickinson).

### Inhibition of EGFR-mediated cell proliferation and signaling

A431 cells were seeded in complete DMEM in 96-well plates. After 24 hours, the medium was replaced by DMEM 1% FBS containing equimolar concentrations (0.19–50 nM) of ctx, atz, IgTT-E or IgTT-1E and incubated for 72 hours. Viability was

assessed using the CellTiter-Glo luminescent assay (Promega, cat# G7570). For EGFR signaling studies, A431 cells were starved overnight in DMEM 1% FBS and then incubated for 4 hours in serum-free DMEM in the presence of 0.1  $\mu\text{M}$  ctx, atz, IgTT-E or IgTT-1E, followed by 5 min incubation with 25 ng/ml of human EGF (MiltenyiBiotec, cat# 130-093-825). After stimulation, cells were lysed in Laemmli lysis buffer, separated under reducing conditions on 4–12% Tris-glycine gels, transferred to nitrocellulose membrane and incubated with the rabbit anti-human phosphor-EGFR (Tyr1068) mAb (clone D7A5; Cell Signaling Technology, cat# 3777) followed by incubation with an IRDye800CW-conjugated donkey anti-rabbit antibody (LI-COR Biosciences, cat# 925–32213). Simultaneously, anti- $\beta$ -actin mouse mAb (Abcam, cat# ab8226) was added as a loading control, followed by IRDye680RD-conjugated donkey anti-mouse (LI-COR Biosciences, cat# 925–68072). Visualization and quantitative analysis of protein bands were carried out with the Odyssey system (LI-COR Biosciences).

### PD-1/PD-L1 Blockade Bioassay

The PD-1/PD-L1 Bioassay (Promega, cat# J1250) was used following manufacturer's instructions. Briefly,  $2.5 \times 10^4$  PD-L1 aAPC/CHO-K1 cells/well were seeded in 96-well white plates in DMEM 10% FBS and incubated overnight at 37 °C. Then, medium was removed and different final concentrations (400; 66.7; 6.67; 0.667 and 0.0667 nM) of atz, ctx, IgTT-E or IgTT-1E were added in 40  $\mu\text{l}$  RPMI 1% FBS/well. Then,  $1.25 \times 10^5$  Jurkat PD-1 cells/well were added in 40  $\mu\text{l}$  RPMI 1% FBS/well and incubated 6 hours at 37 °C. Then, 80  $\mu\text{l}$  of BioGlo Reagent (Promega, cat# G7941) were added and bioluminescence, as an indicator of activation, was measured in a Tecan Infinite F200 Fluorescence Microplate Reader (Life Sciences, Tecan).

### ADCC reporter bioassay

The ADCC reporter Bioassay (Promega, cat# G7010) was used following manufacturer's instructions. Briefly,  $1.2 \times 10^4$  CHO and CHO<sup>EGFR</sup> cells/well were seeded in 96-well white plates in DMEM 10% FBS and incubated overnight at 37 °C. Then, medium was removed and different final concentrations (400, 66.7; 6.67; 0.667; 0.0667 nM) of atz, ctx or IgTT-1E were added in 25  $\mu\text{l}$  RPMI 1% FBS/well. Then,  $7.5 \times 10^4$  ADCC Jurkat<sup>CD16</sup> effector cells/well were added in 25  $\mu\text{l}$  RPMI 1% FBS and incubated 6 hours at 37 °C. Finally, 75  $\mu\text{l}$ /well of BioGlo Reagent (Promega) were added and bioluminescence, as an indicator of ADCC activity, was measured in a Tecan Infinite F200 Fluorescence Microplate Reader.

### Cytotoxicity assay

Peripheral blood mononuclear cells (PBMC) were isolated from peripheral blood of volunteer healthy donors by density gradient centrifugation using lymphoprep (Axis – Shield, cat# AXS-1114544). All donors provided written informed consent in accordance with the Declaration of Helsinki. For cytotoxicity assay, EGFR<sup>+</sup> and PD-L1<sup>+</sup> MDA-MB231<sup>Luc</sup> and A549 cells

were co-cultured with freshly isolated PBMC at two different effector-to-target (E:T) ratios (5:1 and 10:1) in presence of atz, ctx, IgTT-1E or polyclonal control human IgG (6.67 nM). After 48 hours, cells were stained for 30 minutes at 4 °C with V450-conjugated anti-CD45 mAb (Becton Dickinson, cat# 560367) and 7-AAD (BD Biosciences, cat# 559925) in 50  $\mu\text{l}$  of PBS 2% FBS using TruCount Absolute Counting Tubes (BD Biosciences, cat# 663028). Finally, the samples were diluted by adding 450  $\mu\text{l}$  of PBS before proceeding to flow cytometry analysis. Cytotoxicity was determined by recording the residual live target cells (7AAD<sup>-</sup> and CD45<sup>-</sup>).

### NK cells degranulation assay

Degranulation of NK cells following *in vitro* stimulation was assessed by a flow cytometry-based assay. CHO, CHO<sup>PD-L1</sup> and CHO<sup>EGFR</sup> cells were seeded in 96-well U-bottom plates at  $5 \times 10^5$  cells/well and co-cultured with  $2.5 \times 10^6$  PBMC/well (5:1 E:T ratio) in the presence of monensin and PE-labeled anti-CD107a mAb (clone H4A3, BD Biosciences, cat# 555801). Then, IgTT-1E, ctx and atz were added to final concentrations of 6.67; 0.667 and 0.0667 nM. As basal degranulation control, PBMC were plated without target cells. A staining control well received neither degranulation stimuli nor the anti-CD107a mAb. After incubating for 4 hours, plates were centrifuged and cells were resuspended in PBS 0.5% FBS and stained with FITC-labeled anti-human CD3 (clone SK7, BD Biosciences, cat# 345763) and APC-labeled anti-human CD56 (clone B159, BD Pharmingen, cat# 555518) mAbs, and analyzed by a FACSCanto II Flow Cytometer (BD Biosciences). NK cells were identified as CD3<sup>-</sup>CD56<sup>+</sup> events with light scatter characteristics of lymphocytes.

### Mice

NOD.Cg-Prkdc<sup>SCID</sup>IL2rg<sup>tm1Wjl</sup>/SzJ (NSG, The Jackson Laboratory) female mice were housed under pathogen-free conditions with daily cycles of 12 hours light/12 hours darkness, and sterilized water and food were available *ad libitum*. All animal procedures conformed to European Union Directive 86/609/EEC and Recommendation 2007/526/EC, enforced in Spanish law under RD 1201/2005. Animal protocols were approved by the respective Ethics Committee of Animal Experimentation of the participant institutions (imas12 and Instituto Investigación Sanitaria Puerta de Hierro-Segovia de Arana); they were performed in strict adherence to the guidelines stated in the International Guiding Principles for Biomedical Research Involving Animals, established by the Council for International Organizations of Medical Sciences. The experimental study protocols were additionally approved by local government (PROEX 166/19). Mice were kept under anesthesia during all manipulations and all efforts were made to minimize suffering.

### Therapeutic studies

MDA-MB-231<sup>Luc</sup> cells were resuspended in 30% (v/v) matrigel (Corning, cat#356231) in PBS. Cells were injected directly into the second left mammary fat pad of NSG female mice ( $2 \times 10^6$ /

mouse), followed by an intraperitoneal (i.p.) injection of freshly isolated human PBMC ( $1 \times 10^7$ /mouse). Tumor growth was evaluated weekly by bioluminescence imaging. Mice were treated every three days with five intraperitoneal injections of PBS, atz, ctx or IgTT-1E (4 mg/kg), or atz/ctx (2 mg/kg) in combination. Mice weights were measured twice a week to monitor toxicity and animals were euthanized at any sign of distress and/or due to 20% of weight loss. For *in vivo* bioluminescence imaging, mice were anesthetized using inhaled isoflurane, injected intraperitoneally with 125 mg/kg D-luciferin (Promega, cat# E1605) dissolved in 200  $\mu$ l of sterile PBS. Animals were imaged 10 minutes after D-luciferin injection using the Bruker In-Vivo Xtreme (Bruker). The photon flux emitted by the luciferase-expressing cells was measured as an average radiance (photons/sec/cm<sup>2</sup>/sr). Imaging analysis was performed using Bruker Molecular Imaging Software (Bruker). At the end of the experiment anesthetized animals were sacrificed and the different organs (liver, spleen, and tumor) were rapidly harvested and processed. A549 cells ( $2 \times 10^6$ /mouse) resuspended in 30% (v/v) matrigel in PBS were implanted subcutaneously (s.c.) into the dorsal space of NSG mice, followed by an intraperitoneal injection of freshly isolated human PBMC ( $1 \times 10^7$  cells/mouse). Tumor growth was monitored by caliper measurements twice a week. Mice were treated every three days with five intraperitoneal injections of PBS, IgTT-1E (4 mg/kg) or atz/ctx (2 mg/kg) in combination. Mice weights were measured as described above. Mice were euthanized when the weight loss was  $\geq 15\%$  compared to baseline, when tumor size reached a diameter of 1.0 cm in any dimension, when tumors ulcerated, or at any sign of distress.

### Immunohistochemistry

Tumors from different treatment groups were collected at different times after implantation, fixed in 10% neutral buffered formalin (Sigma-Aldrich, cat#HT501128) for 48 hours and after extensive washing in PBS, tissues were embedded in paraffin. Four- $\mu$ m-thick sections were incubated with mouse mAbs listed in Table S2 on a Bond<sup>TM</sup> Automated System (Leica Microsystems) according to the manufacturer's instructions. Nuclei were counter stained with Harris' hematoxylin. Whole digital slides were acquired with a slide scanner (AxioScan Z1, Zeiss), and total versus positive cells were automatically quantified (AxioVision 4.6 software package, Zeiss).

### Statistical analysis

All plots were created using GraphPad Prism 9.0, which was also used to perform statistical analysis. In general, the *in vitro* experiments were done in triplicates and values are presented as mean  $\pm$  SD. Significant differences (*P* value) were identified using a two-tailed, unpaired Student's *t* test assuming a normal distribution, or a one-way analysis of variance (ANOVA), adjusted by Dunnett's test for multiple comparisons, as indicated. *P* values are shown in the corresponding figures. Two-way ANOVA was used to analyze experiments that evaluated the interaction of two variables, such as cell type and therapy, following multiple comparison testing using either Dunnett or Tukey, as appropriate. Mean tumor volumes are presented for

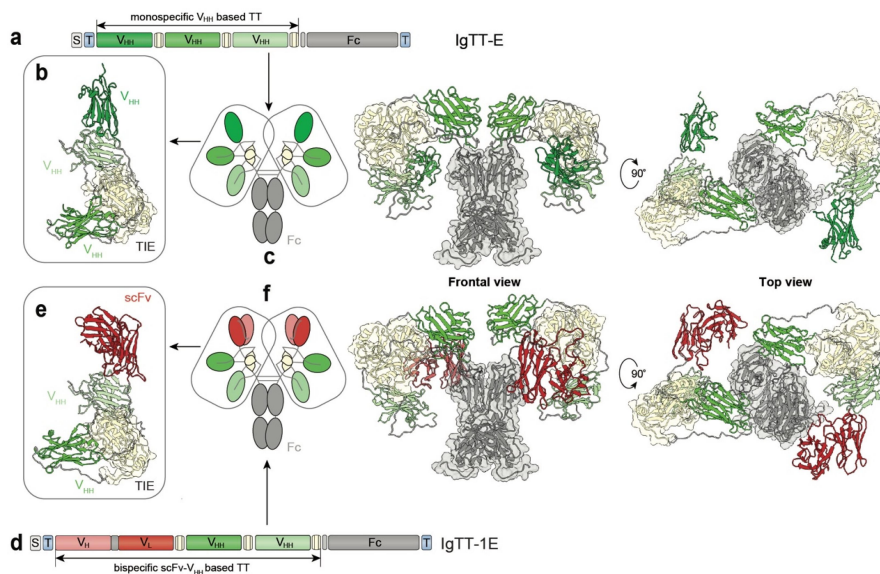
each group using a scatter plot as mean  $\pm$  SD. To assess the differences between treatment groups, *P* values were determined by one-way ANOVA adjusted by the Bonferroni correction for multiple comparison tests. Survival curves were estimated for each group using the Kaplan-Meier method and individual curves were analyzed by Log-rank (Mantel-Cox) test. Significant differences were calculated by comparing each group to the others.

## Results

### Generation and characterization of monospecific and bispecific IgTTs

In this study, we generated an EGFR-specific IgTT (Figure 1a) by fusing a V<sub>HH</sub>-based monospecific TT (Figure 1b), consisting of three identical EGFR-specific V<sub>HH</sub> (clone EGa1)<sup>13</sup>-TIE modules connected by two glycine-serine-based linkers on a single-chain molecule<sup>7</sup>, to the human IgG1 hinge-Fc regions (IgTT-E, Figure 1c). This antibody was further modified (Figure 1d) by replacing the N-terminal V<sub>HH</sub> with an anti-PD-L1 scFv derived from *atezolizumab* (atz) (Figure 1e) to generate a PD-L1/EGFR bispecific IgTT (IgTT-1E, Figure 1f). The models presented in Figure 1 are snapshots of dynamic conformations in which the flexible connectors confer the V<sub>HH</sub> and scFv binding domains variable distances and areas of influence around the Fc region. IgTT antibodies were purified from conditioned medium from stably transfected HEK-293 cells by Strep-Tactin affinity chromatography, and both molecules eluted as single peaks (Fig. a, b). Protein yields for IgTT-E and IgTT-1E were 5.7 and 3.4 mg/L, respectively (Fig. S1a). SEC-MALS measurements for IgTT-E and IgTT-1E resulted in major symmetric peaks with molar masses of 227 and 240 kDa, respectively (Fig. S1a, b), close to the calculated values for each dimeric species in solution (202 and 223 kDa, respectively). In reducing SDS-PAGE, two main bands corresponding to the IgTT-E and IgTT-1E monomers can be identified (Fig. S1c). The smaller minority bands are most likely result of linker cleavage that does not disturb the structural organization of the molecule, as is evident from the observation of single peaks in SEC (Fig. S1a).

The binding of both IgTT antibodies to their cognate antigens was investigated using biolayer interferometry (BLI). Monospecific IgTT-E and bispecific IgTT-1E were found to bind immobilized human EGFR-Fc (EGFR) similarly to *cetuximab* (ctx) (Figure 2a,b), and IgTT-1E bound to human PD-L1-Fc (PD-L1) comparably to atz (Figure 2c). In all four interactions, less than 5% dissociation was measured during a period of 30 minutes, which was not sufficient to accurately determine the dissociation rate constant (KD). The upper bound for their dissociation rate constant is  $3 \times 10^{-5}$ /s and the lower bound for the half-life of their interactions is 7 hours. Therefore, these kinetic experiments demonstrate picomolar affinity or stronger for all the measured interactions, as is expected for bivalent and trivalent antibodies in an experiment which permits multivalent antigen binding, but do not determine precise KD values (Table S3). Additionally, after the binding of IgTT-1E and ctx to immobilized EGFR, the ability of the antibodies to simultaneously bind PD-L1 was investigated by additionally



**Figure 1.** Molecular diagrams and models of the IgTT-E and IgTT-1E antibodies. Gene layout of the monospecific IgTT-E (a), bearing a signal peptide from oncostatin M (white box), three anti-EGFR  $V_{HH}$  (green boxes), three collagen-derived trimerization (TIE) domains (yellow boxes) flanked by peptide linkers and the Fc encoding element (gray boxes). N-terminal FLAG-Strep and C-terminal Myc-His tags (blue boxes) were appended for purification and immunodetection purposes. Schematic diagram showing the three-dimensional model of the TT (b), the molecular diagram and the three-dimensional modelizations of the IgTT-E, in front and top views (c). Gene layout of the bispecific IgTT-1E (d), bearing a signal peptide from oncostatin M (white box), one anti-PD-L1 scFv (red box) and two anti-EGFR  $V_{HH}$  genes (green boxes), three TIE domains (yellow boxes) flanked by peptide linkers and the Fc encoding element (gray boxes). N-terminal FLAG-Strep and C-terminal Myc-His tags (blue boxes). Schematic diagram showing the three-dimensional model of the TT (e), the molecular diagram and the three-dimensional modelizations of the IgTT-1E in front and top views (f).

associating the antibody-loaded biosensors with PD-L1 in solution. IgTT-1E, but not ctx, showed a clear simultaneous binding to PD-L1 while maintaining its interaction with EGFR (Figure 2d), further demonstrating its bispecificity and a lack of steric hindrance between its two cognate interactions. The ability of both antibodies to specifically detect human EGFR and PD-L1 in a cellular context was analyzed by flow cytometry, using ctx and atz as binding controls, and *trastuzumab* (tra) as negative control (Figure 2e). Furthermore, IgTT-1E demonstrated high stability in physiological-like conditions with no significant loss of binding activity after incubation at 37 °C for 96 hours in human serum (Figure S1d).

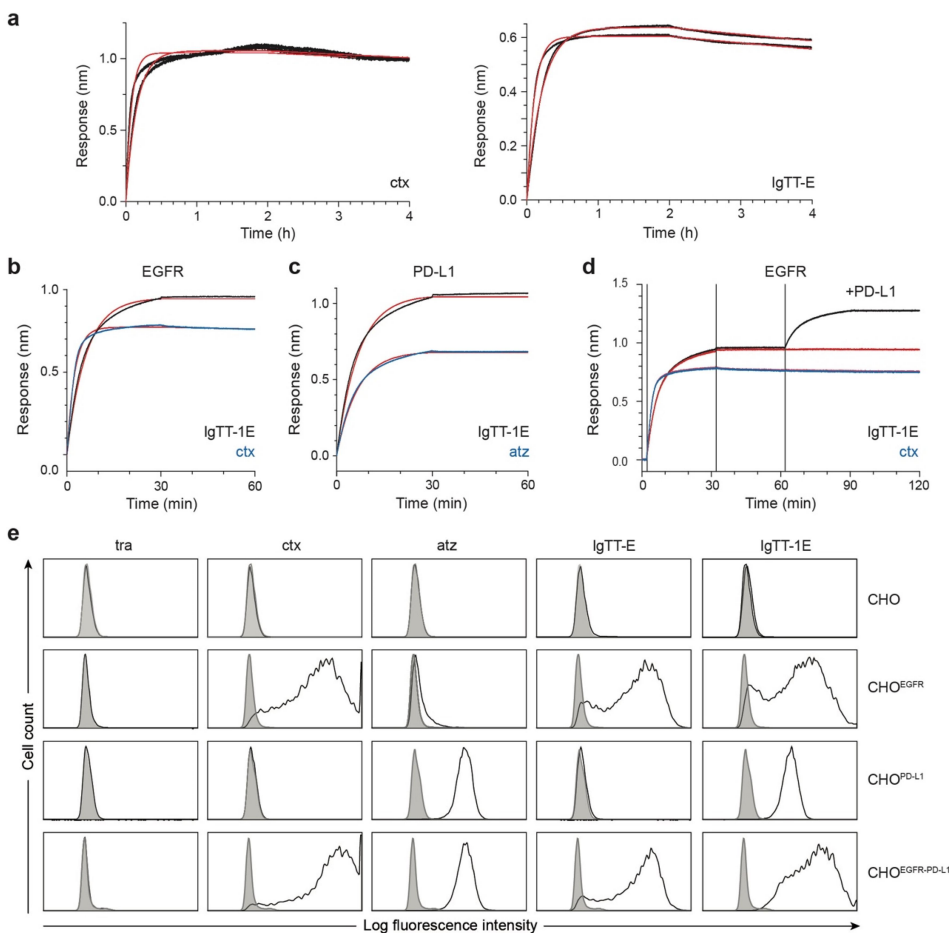
### Effect of IgTT-1E on EGFR-mediated signaling and PD-1/PD-L1 blockade

We next studied the capacity of IgTT-1E to inhibit proliferation and EGFR phosphorylation using the ligand-competitive inhibitor ctx as a positive control and the anti-PD-L1 atz as a negative control. The epidermoid carcinoma A431 cell line is EGFR gene-amplified ( $1.2 \times 10^6$  sites/cell) and mostly depend on the EGFR/MAPK pathway for continuous proliferation<sup>18,19</sup>. The anti-EGFR EgA1  $V_{HH}$  binds an epitope close to junction of EGFR domains II/III, hindering the conformational changes in EGFR necessary for high-affinity ligand binding and receptor dimerization<sup>20</sup>. As shown in Figure 3a, IgTT-E, IgTT-1E and ctx, but not atz, inhibited A431 cell proliferation in a dose-dependent manner ( $P = <0.0001$ , 0.0049 and  $<0.0001$ ; respectively, for the highest doses of both IgTTs, vs. equimolar doses

of control antibody). The phosphorylation status of tyrosine 1068 (Tyr1068) was determined, as this tyrosine is the docking site for Grb2 and its phosphorylation allows the initiation of EGFR mitogenic cascade. Correspondingly, the inhibitory effect on cell proliferation correlated mechanistically with a reduction of EGFR phosphorylation (Figure 3b). The IgTT-1E efficiently blocked PD-1/PD-L1 interaction, as shown by the significant induction of luciferase activity on APC/CHO-K1 cells ( $P < 0.0001$ ), similar to observed with the PD-L1 blocking antibody atz ( $P < 0.0001$ ). In contrast, no PD-1/PD-L1 blocking activity was observed in the presence of ctx or IgTT-E (Figure 3c).

### Determination of Fc-mediated effector functions

For measuring ADCC activity, Jurkat cells constitutively expressing human FcγRIIIa (CD16) on the cell surface and a luciferase reporter driven by a NFAT response element (Jurkat<sup>CD16</sup>) were co-cultured with CHO cells stably expressing human EGFR (CHO<sup>EGFR</sup>) or nontransfected CHO cells as negative control. As shown in Figure 3d, the activation of Jurkat<sup>CD16</sup> cells by ctx and IgTT-1E in co-cultures with CHO<sup>EGFR</sup> cells led to a significant increase in luciferase activity ( $P < 0.0001$  and  $P = 0.0074$ , respectively). In the absence of EGFR-mediated interactions (co-cultures with nontransfected CHO cells), ctx and IgTT-1E showed no induction over untreated Jurkat<sup>CD16</sup> cells (Figure 3d). Similarly, atz showed no activation in co-cultures with CHO<sup>EGFR</sup> or CHO cells (Figure 3d). The atz Fc region has



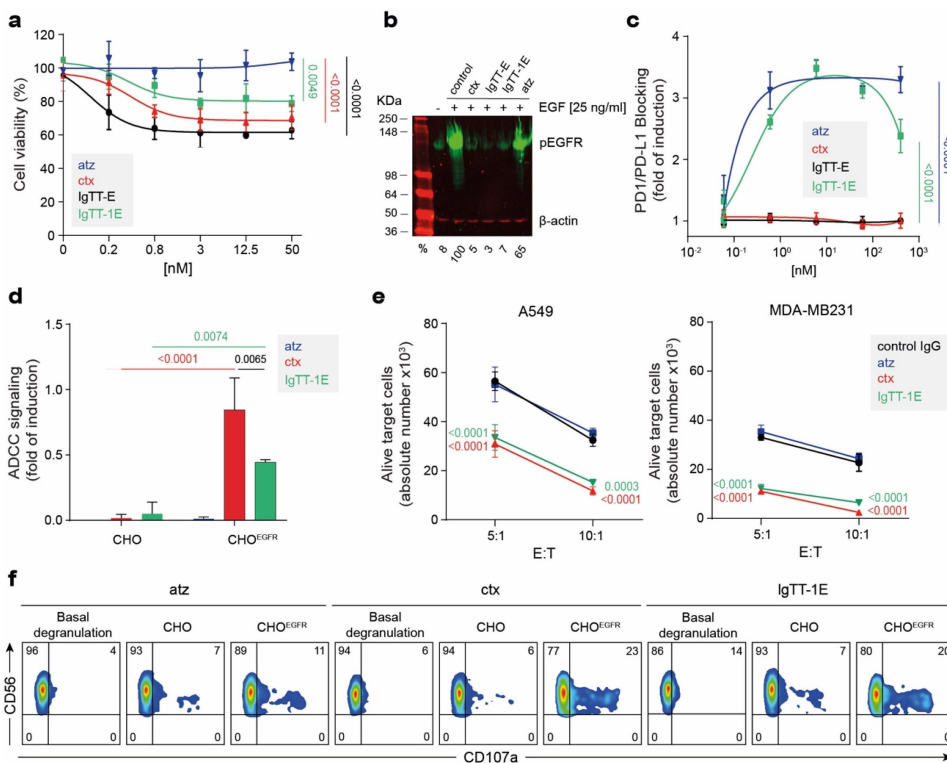
**Figure 2.** Binding characteristics of mono- and bispecific IgTT antibodies to EGFR, PD-L1, and both antigens simultaneously. (a) BLI-derived sensorgrams (in black) for the interaction between cetuximab (ctz) or IgTT-E and immobilized human EGFR-Fc. (b) Human EGFR-Fc was immobilized onto biosensors and 10 nM of either the IgTT-1E (black trace) or ctz (blue trace) was associated for 30 minutes, followed by 30 minutes of dissociation in buffer only. (c) as in the previous panel, binding to immobilized human PD-L1-Fc was investigated using atezolizumab (atz, blue trace) as a comparison antibody. The results of fitting to a 1:1 binding model are shown as red traces. Kinetic rate parameters from fitting are given in Table S3. A single experimental sensorgram of each antibody is shown; duplicate biosensors were included in the experiment and showed negligible variation. (d) Simultaneous binding to both immobilized human EGFR-Fc and human PD-L1-Fc in solution was demonstrated for IgTT-1E (black trace) but not ctz (blue trace). Human EGFR-Fc-coated biosensors were loaded with either IgTT-1E or ctz (as in the panel b), after which biosensors were immersed in 20 nM of human PD-L1-Fc or kept in buffer (red trace). (e) the binding to human EGFR and PD-L1 on the cell surface of CHO, CHO<sup>EGFR</sup>, CHO<sup>PD-L1</sup> and CHO<sup>EGFR-PD-L1</sup> cells by trastuzumab (tra), ctz, atz, IgTT-E and IgTT-1E at 6.67 nM was measured by FACS. Cells incubated with PE-conjugated isotype control mAb are shown as gray-filled histogram. The y-axis shows the relative cell number, and the x-axis represents the intensity of fluorescence expressed on a logarithmic scale.

been engineered to reduce binding to Fcγ receptors and minimize ADCC<sup>21,22</sup>. EGFR<sup>+</sup>PD-L1<sup>+</sup> cancer cells (MDA-MB231 or A549) were co-cultured with human PBMC at two different E:T ratios (5:1 and 10:1) in the presence of atz, ctz, IgTT-1E or control human IgG (6.67 nM). A significant decrease in the residual live target cells was observed with IgTT-1E and ctz compared to atz and control IgG at both E:T ratios. IgTT-1E eliminated nearly 55% of the A549 cells and 70% of the MDA-MD23<sup>Luc</sup> cells at the highest E:T ratio (Figure 3e). Furthermore, the IgTT-1E-mediated degranulation by NK cells was studied after co-culturing CHO and CHO<sup>EGFR</sup> cells with PBMC at a 5:1 E:T ratio in the presence of different antibody concentrations. Both ctz and IgTT-1E increased CD107a cell surface expression

compared to PBMC co-cultured CHO<sup>EGFR</sup> cells in the presence of atz. No changes in CD107a cell expression were detected when PBMC were co-cultured with CHO cells (Figure 3f, Figure. S2 and Table S4).

#### Therapeutic effect of IgTT-1E in vivo

The therapeutic potential of IgTT-1E was first investigated in PBMC-driven humanized NSG mice bearing human MDA-MB-231 TNBC CDXs. MDA-MB-231<sup>Luc</sup> cells were injected into the second right mammary fat pad (MFP) and human PBMC intraperitoneally (i.p.). Mice were treated every three days for five i.p. injections in total, atz, ctz, IgTT-1E or the atz/ctz combination (combo) (Figure 4a). IgTT-1E monotherapy reduced tumor



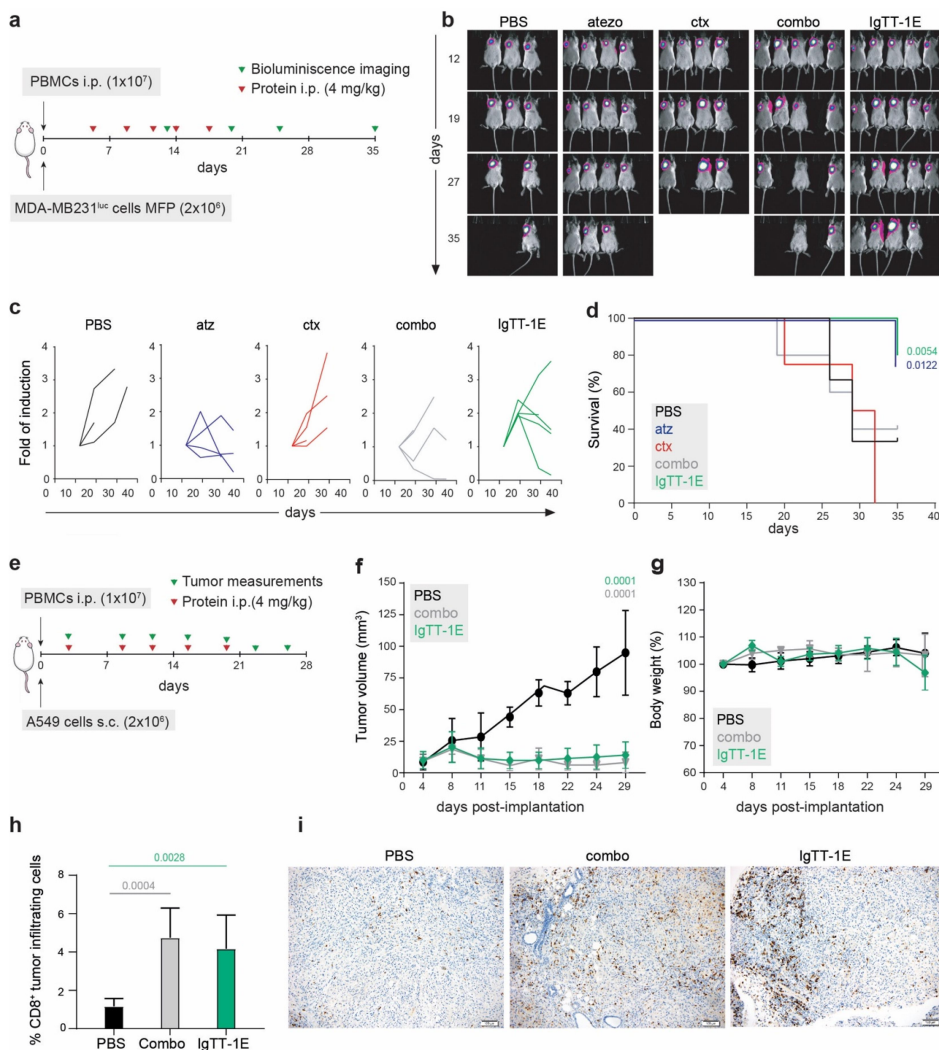
**Figure 3.** IgTT-1E reduces cell proliferation and induces cytotoxicity. (a) Inhibition of A431 cell proliferation through blocking of EGF/EGFR signaling. The cells were treated with the indicated doses of IgTT-E, IgTT-1E, ctx (positive control) or atz (negative control). Viable cells after 72 hours of treatment were plotted relative to untreated controls. Results are expressed as a mean  $\pm$  SD ( $n = 3$ ). Significance was measured by one-way ANOVA with Dunnett's multiple comparisons test. (b) Inhibition of EGFR phosphorylation was assessed by western blotting. After pre-incubation with 50 nM of each antibody for 4 hours, cells were stimulated for 5 minutes with EGF or vehicle. Using the  $\beta$ -actin as load and signal control, the percentages of each phosphorylation band with respect to the positive control were calculated. (c) PD-1/PD-L1 blockade bioassay assesses the inhibitory activities of PD-L1 blocking antibody IgTT-1E. Y-axis represents reporter gene fold induction. Ctx was used as negative control and atz as positive control. Results are expressed as mean  $\pm$  SD ( $n = 3$ ). Significance was measured by one-way ANOVA with Dunnett's multiple comparisons test. (d) ADCC reporter bioassay response to ctx, atz and IgTT-1E, using ADCC bioassay effector Jurkat<sup>CD16</sup> cells against CHO and CHO<sup>EGFR</sup> target cells. Target cells were incubated with ctx (positive control), atz (negative control) and IgTT-1E followed by addition of ADCC bioassay effector cells at a 6:1 E:T ratio. After 6 hours of coincubation at 37°C Bio-Glo™ Luciferase Assay Reagent was added for luminescence determination. Data shown represent the mean  $\pm$  standard deviation of triplicates. Data are presented as the mean  $\pm$  SD ( $n = 3$ ). Quantification shown as mean  $\pm$  SD, ( $n = 3$ ). Significance was determined by two-way ANOVA with Tukey's multiple comparisons test. (e) The number of alive (7AAD<sup>-</sup>) target A549 and MDA-MB231 cells determined after 48 hour co-culture with PBMC cells at 5:1 and 10:1 E:T ratios in presence of control human IgG, atz, ctx and IgTT-1E. Results are shown as mean  $\pm$  SD ( $n = 3$ ). Significance was determined by two-way ANOVA with Dunnett's multiple comparisons test. (f) Human PBMC were incubated with ctx, atz and IgTT-1E (0.667 nM) for 4 hours prior to flow cytometry evaluation of CD107a cell surface expression on NK cells, gated as CD3<sup>+</sup>CD56<sup>+</sup> cells.

bioluminescence in 4 out of 5 treated mice, similar to what was found in the groups receiving atz or combination therapy (Figure 4b,c). Consistent with these antitumor results, a significant overall survival benefit was observed for the mice that had been treated with atz ( $P = 0.0122$ ) or IgTT-1E ( $P = 0.0054$ ) compared to ctx (Figure 4d). To further validate the antitumor activity of IgTT-1E, we performed a second study with PBMC-driven humanized NSG mice bearing human A549 lung cancer CDXs. When tumors reached approximately 0.2 cm in diameter, mice were randomized into groups with similar mean tumor sizes and SDs and treated every three days for five i.p. injections in total of IgTT-1E or the atz/ctx combination (Figure 4e). IgTT-1E monotherapy was able to reduce tumor growth by approximately 85% ( $P = 0.0001$ ), while the

combination therapy showed an approximately 90% tumor growth reduction ( $P = 0.0001$ ) (Figure 4f). There were no significant body weight decreases in any group (Figure 4g). In both treatment groups, inhibition of tumor growth was associated with significantly increased numbers of intratumoral CD8<sup>+</sup> T cells (Figure 4h, i and Figure S3), while no CD56<sup>+</sup> NK cells were detected on tumor tissues (Fig. S4).

## Discussion

In recent years, new cancer immunotherapy strategies have been developed based on the generation of bsAb with improved properties over conventional mAbs<sup>8,12,14,23</sup>. Currently, there are five FDA-approved bsAbs for cancer



**Figure 4.** IgTt-1E treatment promoted significant tumor growth inhibition in NSG mouse models. (a) Timeline of experimental design. MDA-MB-231<sup>Luc</sup> cells were implanted into the right mammary fat pad (MFP) of NSG mice, and human PBMC were injected intraperitoneally (i.p.). Mice were treated with five i.p. injections of PBS, atz, ctx, atz + ctx (combo) or IgTt-1E. (b) Weekly bioluminescence imaging showing tumor progression. (c) Quantification of tumor burden (as bioluminescence fold induction from each mouse) at the indicated time points. (d) Kaplan-Meier plot of survival data. (e) Timeline of experimental design. NSG mice were subcutaneously (s.c.) inoculated with A549 tumor cells and i.p. with human PBMC, follow by five i.p. injections of PBS, atz + ctx combo or IgTt-1E. Tumor progression was monitored weekly by diameter measuring. (f) Average tumor volume growth in each mice group. Data are presented as the mean  $\pm$  SD. Significance was determined by one-way ANOVA adjusted by the Bonferroni correction for multiple comparison test. (g) Change in mouse body weight over time for each group. (h) Quantitative analysis of intratumoral CD8<sup>+</sup> T cells in paraffin-embedded mouse tumor tissue (PBS  $n = 4$ , combo  $n = 6$ , IgTt-1E  $n = 6$ ) by immunohistochemistry. Data were calculated as percentage of CD8<sup>+</sup> versus total cell number and presented as mean  $\pm$  SD. Significance was measured by one-way ANOVA with Dunnett's multiple comparisons test. (i) Immunohistochemical staining for CD8 of representative tissue slides from the tumor of mice treated with PBS, atz + ctx combo or IgTt-1E. Scale bars (100  $\mu$ m) are shown.

treatment, *blinatumomab* for acute B cell lymphoblastic leukemia, *amivantamab-vmjw* for NSCLC, *tebentafusp-tebn* for uveal melanoma, *mosunetuzumab-axgb* for follicular lymphoma, and *teclistamab-cqyv* for multiple myeloma<sup>12,15</sup>, and more than 120 candidates are in clinical development<sup>16</sup>. Despite advances in antibody engineering, the generation of IgG-like bsAb remains challenging when the antigen-binding sites are Fab-like with the variable regions of both the heavy

chain ( $V_H$ ) and light chain ( $V_L$ ), as this makes it difficult to obtain functional bsAb from the random assembly of ten possible  $H_2L_2$  combinations (this is commonly referred to as the chain-association issue)<sup>17–20</sup>. Therefore, to overcome this and other limitations a wide variety of antibody engineering strategies have been developed in recent years<sup>9,23</sup>. Symmetric bsAbs are generated by the assembly of antibodies with unmodified heavy chain (HC) constant ( $C_H$ ) regions<sup>21,23</sup>. Most

symmetric bsAbs in clinical development are bulky molecules with tetravalent (2 + 2) designs in which extra antibody fragments are fused to homodimeric antibody molecules<sup>22,23</sup>. Otherwise, the generation of asymmetric IgG-like bsAbs continues to be challenging because two critical issues need to be simultaneously addressed: the heterodimerization of two different HC chains and the discrimination between the two L/H chain interactions<sup>21,23,24</sup>. Correct HC heterodimerization is facilitated using engineering strategies such as knobs-into-holes and crossMab technologies<sup>23,25–28</sup>. However, most of these strategies employ multiple mutations within the C<sub>H</sub> domains, which can negatively affect the favorable properties associated with native Fc regions, such as high stability and solubility, as well as increasing their immunogenicity<sup>29</sup>.

Here, we generated a novel PD-L1/EGFR IgG-like bsAb by novel antibody engineering that solves some of the problems associated with bispecific heterodimeric IgG antibody generation, such as the chain-association issue<sup>17–20</sup>. The IgTT format is based on the fusion of a mono or multispecific TT<sup>7</sup> with the human IgG1 hinge and Fc region, generating a hexavalent IgG-like antibody capable of bivalently recognizing up to three different antigens. We initially generated a monospecific IgTT molecule targeting EGFR, a well-characterized TAA that is the one of the most commonly overexpressed membranous oncogenic protein in epithelial cancers<sup>30</sup>. In addition, we have demonstrated the suitability of the IgTT format for dual-targeting strategies combining TAA-recognition with an ICB, in a single molecule. Both IgTT molecules specifically bind to their targeted antigens, and the IgTT-1E displayed simultaneous binding to both antigens. As previously described, the interaction of EGal V<sub>HH</sub> with EGFR inhibited EGFR phosphorylation and cell proliferation<sup>7,11,31</sup>. Furthermore, the bispecific IgTT-1E efficiently blocked PD-1/PD-L1 interaction and promoted efficient EGFR-mediated ADCC.

The combined PD-L1 and EGFR targeting by bsAbs has been shown to enhance PD-L1 blockade selectively in the TME, due to EGFR overexpression in cancer cells, and to reduce potential off-tumor binding to PD-L1-expressing normal cells<sup>32</sup>. Because of its distinctive properties including tetravalent binding to EGFR, bivalent binding to PD-L1, and a molecular mass similar to IgG, the IgTT-1E is expected to demonstrate improved tumor localization compared to previously generated bsAbs. Furthermore, EGFR overexpression and activation promotes PD-L1 expression by cancer cells<sup>33</sup>. Therefore, EGFR is a particularly suitable TAA for the development of next-generation bispecific ICB. Compared to other IgG-like bsAbs simultaneously targeting EGFR and PD-L1<sup>34,35</sup>, the IgTT-1E is based on an IgG1 wild-type homodimeric Fc region retaining ADCC activity as in Li et al.<sup>35</sup>. Indeed, IgTT-1E might induce ADCC activity in the intratumoral space reducing the tumor burden and priming immune responses, as has been shown with *cetuximab*<sup>36</sup>. The concomitant presence of PD-L1 blocking domains in the same spatial location would be essential to overcome immunosuppressive counter-

mechanisms in the TME and fully mobilize the adaptive and innate immunity against tumor cells. In addition, the presence of a functional Fc region in the molecule, as in the case of *cetuximab*, has been shown to induce NK cell-dendritic cell (DC) crosstalk, promoting DC maturation and leading to CD8<sup>+</sup> T cell priming<sup>37–47</sup>. Therefore, IgTT-1E may provide sustained CD8<sup>+</sup> T activity by inducing NK-DC crosstalk and blocking PD-1/PD-L1 interaction, leading to a more prolonged function of tumor-infiltrating lymphocytes.

## Acknowledgments

We thank Carlos Alfonso for help with SEC-MALS experiments.

## Disclosure statement

M.C., R.N., and O.H. are all employees of Leadartis. L.A.-V. and L.S. are co-founders of Leadartis. The remaining authors declare no competing interests.

## Funding

L.A.-V. was supported by grants from the MCIN/AEI/10.13039/501100011033 (PID2020-117323RB-I00 and PDC2021-121711-I00), the Instituto de Salud Carlos III (DTS20/00089), the CRIS Cancer Foundation (FCRIS-2021-0090), the Spanish Association Against Cancer (PROYE19084ALVA), the Fundación “La Caixa” (HR21-00761 project IL7R\_LungCan) and the Fundación de Investigación Biomédica 12 de Octubre Programa Investiga (2022-0082). B.B. and L.S. were supported by grants PI20/01030 and PI19/00132 from the Instituto de Salud Carlos III (PI20/01030). FJB and MF-G were supported by grants PID2020-113225GB-I00 and PRE2018-085788 funded by MCIN/AEI/10.13039/501100011033. L.R.-P. was supported by a predoctoral fellowship from the Immunology Chair, Universidad Francisco de Vitoria/Merck. C. D.-A. was supported by a predoctoral fellowship from the MCIN/AEI/10.13039/501100011033 (PRE2018-083445). L.D.-A. was supported by a Rio Hortega fellowship from the Instituto de Salud Carlos III (CM20/00004). O.H. was supported by an industrial PhD fellowship from the Comunidad de Madrid (IND2020/BMD-17668). AE-L was supported industrial PhD fellowship from the Instituto de Salud Carlos III (IFI18/00045).

## ORCID

Laura Rubio-Pérez  <http://orcid.org/0000-0002-2877-6092>  
 Rodrigo Lázaro-Gorines  <http://orcid.org/0000-0002-6885-3486>  
 Seandean L. Harwood  <http://orcid.org/0000-0003-4654-8832>  
 Marta Compte  <http://orcid.org/0000-0002-7138-9266>  
 Rocío Navarro  <http://orcid.org/0000-0002-0083-7711>  
 Antonio Tapia-Galisteo  <http://orcid.org/0000-0002-0507-8435>  
 Jaume Bonet  <http://orcid.org/0000-0001-5210-4387>  
 Belén Blanco  <http://orcid.org/0000-0001-5085-7756>  
 Simon Lykemark  <http://orcid.org/0000-0001-8920-6926>  
 Ángel Ramírez-Fernández  <http://orcid.org/0000-0002-3265-6878>  
 Mariola Ferreras-Gutiérrez  <http://orcid.org/0000-0003-4421-3158>  
 Carmen Domínguez-Alonso  <http://orcid.org/0000-0002-0446-9629>  
 Laura Díez-Alonso  <http://orcid.org/0000-0002-9545-6910>  
 Alejandro Segura-Tudela  <http://orcid.org/0000-0002-5506-0153>  
 Francisco J. Blanco  <http://orcid.org/0000-0003-2545-4319>  
 Cruz Santos  <http://orcid.org/0000-0001-5164-5050>  
 José L. Rodríguez-Peralto  <http://orcid.org/0000-0002-6578-7153>  
 Laura Sanz  <http://orcid.org/0000-0002-3119-3218>  
 Luis Álvarez-Vallina  <http://orcid.org/0000-0003-3053-6757>



## Ethics approval

All procedures involving animals were in accordance with the ethical standards of the corresponding institutional and regional/national committees.

## Data availability statement

The datasets used and/or analyzed during the current study are available from the corresponding author [LA-V] on reasonable request.

## Consent to participate

Informed consent was obtained from all participants included in the study.

## References

- Mahoney KM, Rennert PD, Freeman GJ. Combination cancer immunotherapy and new immunomodulatory targets. *Nat Rev Drug Discov.* 2015;14:561–584. doi:10.1038/nrd4591.
- Sharpe AH, Pauken KE. The diverse functions of the PD1 inhibitory pathway. *Nat Rev Immunol.* 2018;18:153–167. doi:10.1038/nri.2017.108.
- Freeman GJ, Long AJ, Iwai Y, Bourque K, Chernova T, Nishimura H, Fitz LJ, Malenkovich N, Okazaki T, Byrne MC, et al. Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J Exp Med.* 2000;192:1027–1034. doi:10.1084/jem.192.7.1027.
- Wei SC, Duffy CR, Allison JP. Fundamental mechanisms of immune checkpoint blockade therapy. *Cancer Discov.* 2018;8:1069–1086. doi:10.1158/2159-8290.CD-18-0367.
- Morad G, Helmink BA, Sharma P, Wargo JA. Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. *Cell.* 2021;184:5309–5337. doi:10.1016/j.cell.2021.09.020.
- Kubli SP, Berger T, Araujo DV, Siu LL, Mak TW. Beyond immune checkpoint blockade: emerging immunological strategies. *Nat Rev Drug Discov.* 2021;20:899–919. doi:10.1038/s41573-021-00155-y.
- Roskoski R. The ErbB/HER family of protein-tyrosine kinases and cancer. *Pharmacol Res.* 2014;79:34–74. doi:10.1016/j.phrs.2013.11.002.
- Levantini E, Maroni G, Del Re M, Tenen DG. EGFR signaling pathway as therapeutic target in human cancers. *Semin Cancer Biol.* 2022;85:253–275. doi:10.1016/j.semcancer.2022.04.002.
- Antibody therapeutics approved or in regulatory review in the EU or US [Internet]. The Antibody Society. [accessed 2023 March 1]. <https://www.antibodysociety.org/resources/approved-antibodies/>
- Cuesta AM, Sainz-Pastor N, Bonet J, Oliva B, Alvarez-Vallina L. Multivalent antibodies: when design surpasses evolution. *Trends Biotechnol.* 2010;28:355–362. doi:10.1016/j.tibtech.2010.03.007.
- Cuesta AM, Sánchez-Martín D, Blanco-Toribio A, Villate M, Enciso-Álvarez K, Álvarez-Cienfuegos A, Sainz-Pastor N, Sanz L, Blanco FJ, Álvarez-Vallina L. Improved stability of multivalent antibodies containing the human collagen XV trimerization domain. *MAbs.* 2012;4:226–232. doi:10.4161/mabs.4.2.19140.
- Lazarovici P, Lelkes PI. Pardaxin induces exocytosis in bovine adrenal medullary chromaffin cells independent of calcium. *J Pharmacol Exp Ther.* 1992;263:1317–1326.
- Álvarez-Cienfuegos A, Nuñez-Prado N, Compte M, Cuesta AM, Blanco-Toribio A, Harwood SL, Villate M, Merino N, Bonet J, Navarro R, et al. Intramolecular trimerization, a novel strategy for making multispecific antibodies with controlled orientation of the antigen binding domains. *Sci Rep.* 2016;6:28643. doi:10.1038/srep28643.
- Saphire EO, Parren PW, Pantophlet R, Zwick MB, Morris GM, Rudd PM, Dwek RA, Stanfield RL, Burton DR, Wilson IA. Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. *Science.* 2001;293:1155–1159. doi:10.1126/science.1061692.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res.* 2000;28:235–242. doi:10.1093/nar/28.1.235.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215:403–410. doi:10.1016/S0022-2836(05)80360-2.
- Liu S, Miersch S, Li P, Bai B, Liu C, Qin W, Su J, Huang H, Pan J, Sidhu SS, et al. A synthetic human antibody antagonizes IL-18R $\beta$  signaling through an allosteric mechanism. *J Mol Biol.* 2020;432:1169–1182. doi:10.1016/j.jmb.2020.01.012.
- Janmaat ML, Kruyt FAE, Rodriguez JA, Giaccone G. Response to epidermal growth factor receptor inhibitors in non-small cell lung cancer cells: limited antiproliferative effects and absence of apoptosis associated with persistent activity of extracellular signal-regulated kinase or Akt kinase pathways. *Clin Cancer Res.* 2003;9:2316–2326.
- Meira DD, Nóbrega I, de Almeida VH, Mororó JS, Cardoso AM, Silva RLA, Albano RM, Ferreira CG. Different antiproliferative effects of matuzumab and cetuximab in A431 cells are associated with persistent activity of the MAPK pathway. *Eur J Cancer.* 2009;45:1265–1273. doi:10.1016/j.ejca.2008.12.012.
- Schmitz KR, Bagchi A, Roovers RC, van Bergen En Henegouwen PMP, Ferguson KM. Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. *Structure.* 2013;21:1214–1224. doi:10.1016/j.str.2013.05.008.
- Deng R, Bumbaca D, Pastuskovas CV, Boswell CA, West D, Cowan KJ, Chiu H, McBride J, Johnson C, Xin Y, et al. Preclinical pharmacokinetics, pharmacodynamics, tissue distribution, and tumor penetration of anti-PD-L1 monoclonal antibody, an immune checkpoint inhibitor. *MAbs.* 2016;8:593–603. doi:10.1080/19420862.2015.1136043.
- Shah NJ, Kelly WJ, Liu SV, Choquette K, Spira A. Product review on the Anti-PD-L1 antibody atezolizumab. *Human Vacc Immunother.* 2018;14:269–276. doi:10.1080/21645515.2017.1403694.
- Harwood SL, Alvarez-Cienfuegos A, Nuñez-Prado N, Compte M, Hernández-Pérez S, Merino N, Bonet J, Navarro R, Van Bergen En Henegouwen PMP, Lykkemark S, et al. ATTACK, a novel bispecific T cell-recruiting antibody with trivalent EGFR binding and monovalent CD3 binding for cancer immunotherapy. *Oncoimmunology.* 2017;7:e1377874. doi:10.1080/2162402X.2017.1377874.
- Surowka M, Schaefer W, Klein C. Ten Years in the making: application of CrossMab technology for the development of therapeutic bispecific antibodies and antibody fusion proteins. *MAbs.* 2021;13:1967714. doi:10.1080/19420862.2021.1967714.
- Blanco B, Domínguez-Alonso C, Alvarez-Vallina L. Bispecific immunomodulatory antibodies for cancer immunotherapy. *Clin Cancer Res.* 2021;27:5457–5464. doi:10.1158/1078-0432.CCR-20-3770.
- Esfandiari A, Cassidy S, Webster RM. Bispecific antibodies in oncology. *Nat Rev Drug Discov.* 2022;21:411–412. doi:10.1038/d41573-022-00040-2.
- Henricks LM, Schellens JHM, Huitema ADR, Beijnen JH. The use of combinations of monoclonal antibodies in clinical oncology. *Cancer Treat Rev.* 2015;41:859–867. doi:10.1016/j.ctrv.2015.10.008.
- Progress in overcoming the chain association issue in bispecific heterodimeric IgG antibodies - PubMed [Internet]. [accessed 2022 December 19]. <https://pubmed.ncbi.nlm.nih.gov/22925968/>
- Rath T, Baker K, Dumont JA, Peters RT, Jiang H, Qiao S-W, Lencer WI, Pierce GF, Blumberg RS. Fc-fusion proteins and FcRn: structural insights for longer-lasting and more effective therapeutics. *Crit Rev Biotechnol.* 2015;35:235–254. doi:10.3109/07388551.2013.834293.
- Clarke SC, Ma B, Trinklein ND, Schellenberger U, Osborn MJ, Ouisse L-H, Boudreau A, Davison LM, Harris KE, Ugumraj HS, et al. Multispecific antibody development platform based on

- human heavy chain antibodies. *Front Immunol.* 2019;9:3037. doi:10.3389/fimmu.2018.03037.
31. Labrijn AF, Meesters JL, de Goeij BECG, van den Bremer ETJ, Neijssen J, van Kampen MD, Strumane K, Verploegen S, Kundu A, Gramer MJ, et al. Efficient generation of stable bispecific IgG1 by controlled Fab-arm exchange. *Proc Natl Acad Sci U S A.* 2013;110:5145–5150. doi:10.1073/pnas.1220145110.
  32. Wu Y, Yi M, Zhu S, Wang H, Wu K. Recent advances and challenges of bispecific antibodies in solid tumors. *Exp Hematol Oncol.* 2021;10:56. doi:10.1186/s40164-021-00250-1.
  33. Brinkmann U, Kontermann RE. The making of bispecific antibodies. *MAbs.* 2017;9:182–212. doi:10.1080/19420862.2016.1268307.
  34. Marvin JS, Zhu Z. Recombinant approaches to IgG-like bispecific antibodies. *Acta Pharmacol Sin.* 2005;26:649–658. doi:10.1111/j.1745-7254.2005.00119.x.
  35. Merchant AM, Zhu Z, Yuan JQ, Goddard A, Adams CW, Presta LG, Carter P. An efficient route to human bispecific IgG. *Nat Biotechnol.* 1998;16:677–681. doi:10.1038/nbt0798-677.
  36. Labrijn AF, Janmaat ML, Reichert JM, Parren PWI. Bispecific antibodies: a mechanistic review of the pipeline. *Nat Rev Drug Discov.* 2019;18:585–608. doi:10.1038/s41573-019-0028-1.
  37. Spiess C, Zhai Q, Carter PJ. Alternative molecular formats and therapeutic applications for bispecific antibodies. *Mol Immunol.* 2015;67:95–106. doi:10.1016/j.molimm.2015.01.003.
  38. J-H H, Kim J-E, Kim Y-S. Immunoglobulin Fc heterodimer platform technology: from design to applications in therapeutic antibodies and proteins. *Front Immunol.* 2016;7:394. doi:10.3389/fimmu.2016.00394.
  39. Ridgway JB, Presta LG, Carter P. 'Knobs-into-holes' engineering of antibody CH3 domains for heavy chain heterodimerization. *Protein Eng.* 1996;9:617–621. doi:10.1093/protein/9.7.617.
  40. Nuñez-Prado N, Compte M, Harwood S, Álvarez-Méndez A, Lykkemark S, Sanz L, Álvarez-Vallina L. The coming of age of engineered multivalent antibodies. *Drug Discov Today.* 2015;20:588–594. doi:10.1016/j.drudis.2015.02.013.
  41. Thomas R, Weihua Z. Rethink of EGFR in cancer with its kinase independent function on board. *Front Oncol.* 2019;9:800. doi:10.3389/fonc.2019.00800.
  42. Akbay EA, Koyama S, Carretero J, Altabel A, Tchaicha JH, Christensen CL, Mikse OR, Cherniack AD, Beauchamp EM, Pugh TJ, et al. Activation of the PD-1 pathway contributes to immune escape in EGFR-driven lung tumors. *Cancer Discov.* 2013;3:1355–1363. doi:10.1158/2159-8290.CD-13-0310.
  43. Chen N, Fang W, Zhan J, Hong S, Tang Y, Kang S, Zhang Y, He X, Zhou T, Qin T, et al. Upregulation of PD-L1 by EGFR activation mediates the immune escape in EGFR-Driven NSCLC: implication for optional immune targeted therapy for NSCLC patients with EGFR mutation. *J Thorac Oncol.* 2015;10:910–923. doi:10.1097/JTO.0000000000000500.
  44. Koopmans I, Hendriks D, Samplonius DF, Ginkel RJ, Heskamp S, Wierstra PJ, Bremer E, Helfrich W. A novel bispecific antibody for EGFR-directed blockade of the PD-1/PD-L1 immune checkpoint. *Oncoimmunology.* 2018;7:e1466016. doi:10.1080/2162402X.2018.1466016.
  45. Li L, Deng L, Meng X, Gu C, Meng L, Li K, Zhang X, Meng Y, Xu W, Zhao L, et al. Tumor-targeting anti-EGFR x anti-PD1 bispecific antibody inhibits EGFR-overexpressing tumor growth by combining EGFR blockade and immune activation with direct tumor cell killing. *Transl Oncol.* 2021;14:100916. doi:10.1016/j.tranon.2020.100916.
  46. Ferris RL, Lenz H-J, Trotta AM, Garcia-Foncillas J, Schulten J, Audhuy F, Merlano M, Milano G. Rationale for combination of therapeutic antibodies targeting tumor cells and immune checkpoint receptors: harnessing innate and adaptive immunity through IgG1 isotype immune effector stimulation. *Cancer Treat Rev.* 2018;63:48–60. doi:10.1016/j.ctrv.2017.11.008.
  47. Srivastava RM, Lee SC, Andrade Filho PA, Lord CA, Jie H-B, Davidson HC, López-Albaitero A, Gibson SP, Gooding WE, Ferrone S, et al. Cetuximab-activated natural killer and dendritic cells collaborate to trigger tumor antigen-specific T-cell immunity in head and neck cancer patients. *Clin Cancer Res.* 2013;19:1858–1872. doi:10.1158/1078-0432.CCR-12-2426.



The thesis was conducted within the Biomolecular NMR group at  
**Center for Biological Research (CIB-CSIC)**  
and  
**CICbioGUNE**  
under the supervision of  
**Dr. Francisco José Blanco Gutiérrez**

**CICbioGUNE**  
MEMBER OF BASQUE RESEARCH  
& TECHNOLOGY ALLIANCE

