Osteocalcin binds to the GPRC6A Venus fly trap allosteric site to positively modulate GPRC6A signaling

Rupesh Agarwal*[1,2], Min Pi*[3], Ruisong Ye[3], Micholas Dean Smith[1,2], Jeremy C. Smith[1, 2], L. Darryl Quarles[3]

*Authors have contributed equally

[1]University of Tennessee/Oak Ridge National Laboratory Center for Molecular Biophysics, Oak Ridge, Tennessee 37830; and [2]Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996. [3]Departments of Medicine, University of Tennessee Health Science Center, Memphis, Tennessee 38163.

Running title (60 characters): Osteocalcin binds to the GPRC6A as a positively allosteric modulator

Keywords: G protein-coupled receptor (GPCR), homology modeling, mutagenesis, Osteocalcin, isoform-selectivity, positively allosteric modulator, extracellular domain

Abstract

GPRC6A is a member of the Family C G-protein coupled receptors that is activated by cations, L-amino-acids, osteocalcin (Ocn) peptide, and testosterone and functions as a master regulator of energy metabolism and sexual reproduction. Based on homology to related receptors, mGlur5 and CaSR, GPRC6A's multiple ligand specificity is likely based on an orthosteric ligand binding domain in the bilobed venus fly trap (VFT) and two positive allosteric modulator (PAM) sites, one in the VFT and the other in the 7TM domain. In these studies, we show that Ocn acts as a PAM for GPRC6A, with binding to a site in the VFT distinct from the orthosteric site for calcium and amino-acids. In agreement with this model, alternatively spliced GPRC6A isoforms 2 and 3, which lack regions of the VFT, and mutations in the predicted Ocn binding site, K352E and H355P, prevented Ocn activation of GPRC6A. These observations provide a structural framework for understanding the ability of multiple distinct classes of compounds to activate GPRC6A and set the stage to develop novel small molecules to activate and inhibit this receptor.

Introduction

Osteocalcin (Ocn) is a 49 amino acid protein in humans that is mainly produced by osteoblasts in bone. A vitamin K-dependent carboxylated form of Ocn comprises much of the non-collagenous bone protein matrix, while an undercarboxylated form (unOcn) is released into the circulation to function as a hormone. unOcn regulates energy metabolism through direct effects on metabolic functions in multiple target tissues and by indirect effects of an ensemble of metabolically active hormones whose release is stimulated by unOcn (1-5). Indeed, unOcn is reported to stimulate insulin secretion and β-cell proliferation in the pancreas (5), fibroblast growth factor 21 (FGF-21) release and glucose and fat metabolism in liver hepatocytes (6), interleukin 6 (IL-6) secretion and glucose utilization in skeletal muscle myocytes (7), adiponectin release and lipolytic activity in adipocytes in whit fat (8), testosterone production from testicular Leydig (2,6), and glucagon-like peptide 1 (GLP-1) secretion from gastrointestinal enterocytes (9,10).

The peripheral metabolic effects of Ocn are purported to be mediated through activation of GPRC6A, a member of the Family C G-protein coupled receptors (5,11). GPRC6A has been previously shown to be activated by several orthosteric ligands, including basic L-amino acids (such as L-Lys, L-Arg, and L-ornithine) and divalent cations (such as Ca^{2+} and Mg^{2+}). In addition, positive allosteric modulators (PAMs) of GPRC6A include testosterone, NPS-568, di-phenyl and tri-phenyl compounds (12,13). Natural products in green tea, gallate and epigallocatechin 3-gallate (ECGC), respectively act to activate and inhibit GPRC6A (14). The structural basis for GPRC6A interaction with multiple ligands is incompletely understood.

Family C GPCRs are characterized by both their similarity in the sequences of their 7-helicies containing transmembrane domain, expression as functional dimers and, in most cases, the existence of large N-terminal endogenous ligand-binding Venus flytrap domains (15). Recently, high resolution structural data of the related mGluR5 (16) and CaSR (17-20), have identified how various ligands and drug molecules target the VFT and 7TM. For example, mGluR5 has the orthosteric binding site for L-glutamate in the VFT, and two positive modulating allosteric (PAM) sites, one located in the VFT and the other in the 7TM (16). The allosteric site located in the VFT was discovered by a nanobody, Nb43, that binds to the Helix L and L-M loop of mGluR5 to potentiate orthosteric agonist binding and receptor activation. Similarly, elucidation of the 3D structure of CaSR identified the orthosteric binding site for Ca^{2+} ion that acts as a composite agonist with L-amino acids, such as L-tryptophan, to stabilize the closure of active VFT (17) and an allosteric site at the dimeric interface of the VFT (proximal to residue D248) where the peptide PAM etecalcitide binds. In CaSR, additional small molecule drugs act as positive allosteric modulators

(PAMs) (e.g., calcimimetics), or negative allosteric modulators (NAMs) (e.g., calcilytic) via a site in the 7TM (17-20).

Based on this new structural information, we developed a structural model for the activation of GPRC6A by orthosteric and allosteric ligands. We investigated Ocn activation of GPRC6A isoforms with deleted segments of the extracellular domain, mapped the peptide fragments of Ocn necessary for activation of GPRC6A, and examined if Ocn peptides bind to the VFT and modulate GPRC6A function.

Results

Structural modeling of GPRC6A isoforms and L-Arginine orthosteric ligand binding.

GPRC6A exists as three isoforms, isoform 1 represents the full-length receptor, and two alternatively spliced receptor isoforms 2 and 3 are characterized by respective deletions of exon 3 and exon 4 segments that encode regions of the VFT domain (21) (**Figure 1A**). To identify regions that differ between the isoforms, we aligned the protein sequence of isoforms 1, 2 and 3 of GPRC6A (**Figure 1B, S1**). The TM domains of these three isoforms have 100% sequence identity, but isoforms 2 and 3 are missing respective residues 271-445 and 446-516 in the extracellular domain.

To define the structure of GPRC6A, we used AlphaFold2 (22-25) to model the full-length isoform 1 **(Figure 1c)**. We confirmed that GPRC6A has 3 domains: TM, CRD and VFT similar to the related mGluR5 and CaSR cryo-EM structures. The residues deleted in the isoforms 2 and 3 correspond to different segments of the VFT domain. In case of isoform 2, a large segment of the VFT domain structure (~ 174 residues) is deleted, whereas a comparatively small region (~ 70 residues) is missing in isoform 3 (**Figure S1**).

In isoform 1, an orthosteric ligand site near residues Tyr 148, Ser 149, Thr 172, Asp 303 was identified in the VFT based on structure superimposition to mGlurR5 and CaSR (**Figure S2a and S2b**). In contrast, in isoform 2, the orthosteric ligand binding site predicted based on aligning orthosteric bound structures of mGluR5 and CaSR structures is missing, foretelling a loss-of-function (*vide infra*) (**Figure S2c)**. The predicted orthosteric ligand binding site is retained in isoform 3, and hence this isoform is predicted to be responsive to orthosteric ligands (*vide infra*).

To test these predictions, GPRC6A isoforms 1, 2 or 3 were transfected into HEK-293 cells and receptor activation measured by cAMP accumulation in response to the orthosteric ligand L-Arg

96 in the presence of 1 mM calcium. Isoform, 2 resulted in loss of L-Arg stimulation of cAMP,
97 whereas isoform 3 retained responsive to L-Arg activation, similar to isoform 1 (**Figure 2A**).

98 ***GPRC6A isoforms 2 and 3 have attenuated responses to Ocn.***

99 The location of Ocn binding GPRC6A is still not well understood. To explore the role of the VFT
100 in mediating the response to Ocn, we compared the ability of Ocn to activate the three GPRC6A
101 isoforms (**Figure 2B and C**). Ocn in the presence of 1 mM calcium activated GPRC6A isoform 1
102 in a dose dependent manner, consistent with previous reports (5). GPRC6A isoform 2 exhibited
103 an attenuated ERK response but a complete loss of cAMP activation in response to Ocn treat-
104 ment. In contrast, GPRC6A isoform 3 showed attenuation of Ocn-induced ERK stimulation but
105 the cAMP response to Ocn persisted (**Figure 2B and C**). Taken together, loss-of Ocn responses
106 in isoforms 2 and 3 suggest that the VFT is necessary for Ocn activation of GPRC6A and that the
107 different isoforms bias ligand-induced coupling to G-protein signaling.

108 ***Identifying Ocn amino acids necessary for GPRC6A activation***

109 Ocn in humans is a 49 amino acid peptide. Initially, we used functional assays to refine the
110 protein-protein interface between Ocn and GPRC6A VFT. We designed Ocn fragments (**Figures**
111 **3A**) (11,26) and compared their ability to activate the full length GPRC6A isoform 1 (**Figure 3B**
112 **and C**) transfected into HEK293 cells. Ocn peptide activation was measured by assessing both
113 ERK (**Figure 3B**) and cAMP (**Figure 3C**) signaling. Compared to the full length Ocn (1-49), all
114 peptide fragments activated signaling in cell transfected with GPRC6A to some degree. However,
115 we found that the decapeptide fragment of Ocn 20-29 (revcelnpdc) located in the middle of the
116 protein and the C-terminal hexapeptide Ocn 44-49 (rfygpv) exhibited a significantly higher activity
117 (**Figure 3B and C**).

118 ***Protein-protein docking to generate GPRC6A-VFT: Ocn complex structure***

119 As noted above, there are two PAM sites in the related receptors mGluR5 and CaSR, one in the
120 7TM domain targeted by small molecules and another in the VFT targeted by the protein nano-
121 body (16) and the etecalcitide peptide (20). Specifically, the PAM function of Nb43 is mediated by
122 a nanobody binding site near a helix in the VFT domain of the mGlur5 as defined by Cryo-EM
123 (16). Aligning the structures of the allosteric ligand bound VFTs mGluR5 (PDB 6N4Y) and CaSR
124 (PDB 7E6U, 7M3G) with our model of GPRC6A and its isoforms indicated that the dimeric inter-
125 face found to bind etecalcitide and NAM in CaSR was invariant across the GPRC6A isoforms,
126 while the binding pocket homologous to the mGluR5 nanobody PAM site was absent in
127 GPRC6A's isoform 2 (**Figure S3a**). We hypothesized that, given the loss of Ocn associated

128 function in iso2, the pocket homologous to the mGluR5 binding site may serve as binding site for
129 Ocn in GPRC6A. Using this site as our guide, we used HADDOCK 2.4 (27) to generate a model
130 of the GPRC6A-VFT: Ocn complex. The top 4 scoring structures within the calculated top cluster
131 were then subjected to additional analysis and visual inspection.

132 Our analysis revealed that the GPRC6A-VFT forms multiple hydrogen bonds (a list is provided in
133 **Table 1)** with Ocn in all the models. Residues 20-29 from highest active Ocn fragment also forms
134 multiple hydrogen bonds with GPRC6A residues (Gln 19, Asp26, Gln381, Lys352) (**Figure 4**) and
135 is also in proximity (within 4 Å) to other residues (Ala385, Arg382, Asp26, Asp349, Gln19, Gln22,
136 Gln61, Glu356, Glu62, His351, His355, Pro20, Ser348, Ser350, Val28), which can form other
137 potential electrostatic interactions. Moreover, the Ocn model shows that there is a disulfide bond
138 between residues Cys 23 and Cys 29 (**Figure S3b**), both of which are present in the Ocn frag-
139 ment. This is predicted to provide structural integrity to the fragment and would potentially fold
140 similarly to the full length Ocn peptide. In analyzing the interactions between the GPRC6A-VFT
141 and Ocn, we observe that two GPRC6A-VFT residues, Lys 352 and His 335, in the helix region
142 (352-363) are common in all the models (**Figure 4, Table 1**). These residues form hydrogen
143 bonds with Ocn residues Asp 28, Asp 30 and Glu 31 and are near several other residues present
144 in 20-29 region of Ocn that was the most active fragment.

145 ***Validation of the complex structure using single point mutagenesis***

146 To test the above modeling, we mutated Lys352, or His355 in human GPRC6A, into Glu352 (mu-
147 tant K352E) or Pro355 (mutant H355K), and also created double mutated Glu352/Pro355 (mutant
148 K352E/H355P) GPRC6A. Mutants and wild-type hGPRC6A cDNAs were transiently transfected
149 into HEK-293 cells and signaling assessed in response to Ocn treatment.

150 We confirmed that mutants K352E, H355P, K352E/H355P, and WT mGRPC6A proteins were
151 equally expressed, as assessed by Western blotting using a Myc antibody, which recognized the
152 Myc epitope located at after signal peptides (1-18 amino acids) of the WT and mutant receptors
153 (**Figure S4 a,b**)

154 The full-length Ocn 1-49 peptide dose-dependently stimulated cAMP accumulation in the wild-
155 type receptor, but the response was significantly decreased in the K352E, H355P, K352E/H355P
156 mutant receptors (Figure 5A). Cells transfected with these three mutants hGPRC6A cDNAs also
157 showed no response to Ocn 20-29 stimulation (**Figure 5B**). These results indicate that the binding
158 sites identified by the computational modeling are accurate and that the helical region (352-363)
159 is important for Ocn binding.

160 The Ocn 44-49 C-terminal fragment, which was predicted to bind to the 7TM in our prior homolo-
161 gous model (11), also had a significantly attenuated signaling response in the single mutants.
162 Interestingly, Ocn 44-49 was able to partially activate the double mutant as assessed by cAMP
163 accumulation (**Figure 5B**).

164

165 ## Discussion

166 Recent cryo-EM analysis of Family C mGluR5 (16) and CaSR (17,19,20) receptors describes the
167 structural basis for orthosteric ligand activation by calcium and L-amino acids and allosteric mod-
168 ulating effects of peptides and small molecules. In the current study, we developed homologous
169 structural models of GPRC6A and found evidence for both orthosteric and allosteric binding sites
170 in its VFT, similar to mGLUR5 and CaSR.

171 GPRC6A is activated by L-amino acids, cations and the Ocn peptide (4). Loss of Ocn activation
172 in isoform 2, however, called into question our prior predictions of the binding of the hexapeptide
173 C-terminal domain of Ocn to the 7-TM (11). In addition**,** comparison of our GPRC6A model with
174 existing cryo-EM structures of mGluR5 and CaSR in the presence of their respective peptide PAM
175 ligands raised the possibility for binding of the Ocn peptide in the VFT domain of GPRC6A. Com-
176 plimentary peptide-protein docking calculations and mutagenesis studies identified an Ocn bind-
177 ing site within the GPRC6A VFT located near the helical region (352-363).

178 Reports of GPRC6A activation by the peptide Ocn are inconsistent, with some studies showing
179 activation (2,3,5,11) and other studies finding no effect (28). The location of Ocn binding site in
180 the VFT at a site distinct from the orthosteric binding site suggests Ocn functions as a PAM rather
181 than an orthosteric ligand. The fact that Ocn activation of GPRC6A requires the presence of
182 calcium (3,5,11) is consistent with Ocn function as a PAM for GPRC6A (2,11).

183 Our findings also suggest altered ligand specificity and downstream signaling of GPRC6A
184 isoforms created by alternative splicing. The loss of a segment of the VFT in isoform 2 resulted in
185 the loss of responses to L-Arg as well as Ocn. In contrast, isoform 3 retains responsiveness to
186 orthosteric activation but exhibits biased signaling to Ocn, as evidence by loss of ERK but preser-
187 vation of cAMP signaling. Loss of this segment may alter the conformational changes leading to
188 opening of binding sites in the intracellular TM 6 domain that is important for G-protein coupling.
189 If so, different tissue distributions or the ratio of these isoforms would be expected to change
190 orthosteric ligand and PAM sensitivities as well as alter coupling to downstream signaling path-
191 ways resulting in tissue-specific regulation mechanisms.

Our modeling data shows that most contacts between Ocn and GPRC6A are driven by the structural part of the Ocn peptide (3 helical structure). Peptide fragments of Ocn consisting of 1-7, 8-43, and 44-49 are released from bone (29). We tested the GPRC6A activation potential of multiple Ocn fragments to further identify the key regions of Ocn that interface with GPRC6A. The hexapeptide 20-29 most strongly activated GPRC6A, consistent with its location in the crucial region of the full-length Ocn used in our docking studies to the VFT helix region. Ocn **8-19 (**APVPYPDPLEPR) also activated GPRC6A (**Figure 3**). This upstream peptide is similar to the pentadecapeptide peptide WLGAPVPYPDPLEPR that is reported to activate GPRC6A *in vitro* and to improve fatty liver disease and insulin resistance after oral or intraperitoneal administration to a mouse model (30). When examining the activities associated with 44-49, it is important to recall that previous docking and mutagenesis studies suggested that the C-terminus of Ocn binds to the extracellular transmembrane domain of GPRC6A (11). This might explain the effects of the RFYGPV, C-terminal peptide of Ocn to activate GPRC6A in setting of mutations to disrupt the Ocn binding domain in the VFT helix. Regardless, short peptides derived from Ocn might be developed for therapeutics to activate GPRC6A, similar to the 7 amino acid etacalcitide PAM for CaSR.

Another interesting observation is the greater potency of the hexapeptide 20-29 compared to Ocn 1-49. Hypothetically this may be due to a potential disulfide bond present in the small peptide that creates a stable complementary conformation within the predicted binding site. Alternatively, competitive self-interactions between the domain and the disordered N-terminus of Ocn may inhibit long-lived binding conformations between Ocn and GPRC6A. Instead, flexibility of the N-terminal domain that result in N-terminal-VFT contacts may inhibit conformational changes along the dimer interface of the VFT domains.

Although we did not examine binding of other PAMs to GPRC6A in this study, testosterone and triphenol compounds that activate GPRC6A are predicted to bind to the 7TM domain of GPRC6A (12,13), analogous to the second PAM domain in the 7TM domain of mGluR5 and CaSR (16,20). This suggests that members of the Family C GPCRs, at least those with large VFT domains, have a conserved structure consisting of two allosteric binding sites, one in the VFT and the other in the 7TM domain, along with an orthosteric ligand binding domain in the VFT.

Finally, GPR158 is another proposed Ocn sensing receptor. The Cryo-EM structure of GPR158 has recently been defined (31). Our VFT Ocn binding region between GPRC6A and 158 is not

224 conserved and a structure superimposition of the two structures shows that the helix region (352-
225 363) identified for GPRC6A is missing in GPR158.

226 Our findings are limited by the lack of Cryo-EM structural data for GPRC6A. Cryo-EM analysis of
227 mGlurr5 and CaSR show that each functions as homodimers, where activation leads to closure
228 of the VFT and conformational changes the 7TMs revealing G-protein binding sites on the cyto-
229 plasmic side necessary for signaling(17). For CaSR, which is most closely related to GPRC6A,
230 receptor activation results in 7TM asymmetry in the homodimer leading to only one protomer for
231 G-protein coupling being stabilized by PAM binding. Our biased signaling response to Ocn in
232 isoforms 2 and 3 suggests that there may be a structural basis for the divergent cAMP and ERK
233 responses to Ocn. To fully understand GPRC6A signaling we may need cryo-electron microscopy
234 structures of GPRC6A in inactive or active states bound to the orthosteric ligands, PAMS and
235 NAMs.

236 Experimental procedures

237 System and docking details

238 GPRC6A and Ocn structures were modeled using AlphaFold2 (22,25,32), which is an AI system
239 developed by DeepMind to predict a 3D structure of a protein from its amino acid sequence.
240 Multiple sequence alignment (MSA) was using Jackhammer method. The 'max_recycles' param-
241 eter,r which controls the maximum number of times the structure is fed back into the neural net-
242 work for refinement, was set to 3. All the other parameters were set to default. The final models
243 were ranked using the pLDDT score. Glycans were not added as the details are not well known.
244 For docking, HADDOCK 2.4 (27) was used to dock the GPRC6A-VFP to Ocn. GPRC6A residue
245 numbers 334,335,336,337 and mature Ocn residue numbers 21,22,23,24,25,26, and 27 were
246 used to guide the docking. The best 4 models from the top cluster (containing 133 models) with
247 HADDOCK score of -81.2±1.1 and Z-score of -1.5 were analyzed.

248 Simulation details

249 Molecular dynamics simulation (MD) of the top ranked AlphaFold2 model was performed using
250 the Amber simulation engine(33) to refine the structures. The MD was performed on the protein
251 using Amber14 ff14SB (34) force field, with a non-bonded cutoff of 10 Å using the Particle Mesh
252 Ewald algorithm (35). The protein-glycan bound system was hydrated by water model TIP3P (36)
253 in an octahedral box of 10 Å around the protein in each direction. Initially, the protein was held
254 fixed with a force constant of 500 kcal mol^{-1} $Å^{-2}$ while the system was energy minimized with 500
255 steps of steepest descent followed by 500 steps with the conjugate gradient method. In a second

256 minimization step, the restraints on the protein were removed and 1000 steps of steepest descent
257 minimization was performed, followed by 1500 steps of conjugate gradient. The system was
258 heated to 300 K while holding the protein fixed with a force constant of 10 kcal mol^{-1} $Å^{-2}$ for 1000
259 steps. Then, the restraints were removed, and 1000 MD steps were performed. The SHAKE al-
260 gorithm(37) was used to constrain all bonds involving hydrogen in the simulations. MD production
261 runs were performed at 300 K using the NPT ensemble and a 2 fs time step. The temperature
262 was fixed with the Langevin dynamics thermostat (38) and the pressure was fixed with the Monte
263 Carlo barostat (39). The last snapshot from 20ns trajectory was used for all analyses.

264 **Measurement of Total and Phospho-ERK by ERK Elisa Analysis**

265 Ocn was purified from bovine tibial bone extracts (40,41). Decarboxylated Ocn was produced by
266 treating Ocn in vacuo at 110°C (41-43). Purity and decarboxylation state were confirmed by native
267 gel electrophoresis (40), or by blotting followed by reaction with 4-diazobenzene sulfonic acid
268 staining for γ-carboxyglutamic acid (41,42). The human Ocn fragments were synthesized by Bio-
269 Matik USA (Wilmington, Delaware, USA).

270 All culture reagents were from Invitrogen (Waltham, MA, USA). Human embryonic kidney HEK-
271 293 cells were obtained from American Type Culture Collection and cultured in DMEM medium
272 supplemented with 10% fetal bovine serum and 1% Penicillin/Streptomycin (P/S). Briefly, HEK-
273 293 cells transfected with/without human GPRC6A isoforms cDNA plasmids were starved by
274 overnight incubation in serum-free DMEM/F12 containing 0.1% bovine serum albumin (BSA) and
275 stimulated with various ligands at different doses. ERK activation were assessed 20 min after
276 treatment by using ERK1/2 (phospho-T203/Y204) ELISA Kit (Invitrogen) corrected for the amount
277 of total ERK using ERK1/2 (Total) ELISA Kit (Invitrogen) to measure ERK levels.

278 **Measurement of cAMP accumulation**

279 HEK-293 transfected with human GPRC6A isoforms cDNA plasmids cells (10^5 cells/well) (44)
280 were cultured in triplicate in 24-well plates in DMEM supplemented with 10% fetal bovine serum
281 and 1% penicillin/streptomycin (100 U/mL of penicillin and 100 μg/mL of streptomycin) for 48
282 hours followed by 4 hours incubation in DMEM/F12 containing 0.1% BSA and 0.5 mM IBMX to
283 achieve quiescence. Quiescent cells were treated with vehicle control, various ligands at concen-
284 tration as indicated for 40 minutes at 37°C. Then, the reaction was stopped, and the cells lysed
285 with 0.5 mL 0.1N HCl. cAMP levels were measured by using Cyclic AMP EIA kit (Cayman Chem-
286 ical) following the manufacture's protocol.

287 **Site-directed mutagenesis**

In-vitro mutagenesis by PCR-mediated recombination QuikChange II XL Site-Directed Mutagenesis Kit (Agilent) was performed using the cDNA of human GPRC6A cloned in the vector pcDNA3 (Invitrogen) as template. The primer sets as following. K352E.sen: GTGACAGTCACGAACTCTTACATG; K352E.antisen: CATGTAAGAGTTCGAGACTGTCAC; H355P.sen: CAAACTCTTACCTGAATATGCCATG; H355P.antisen: CATGGCATATTCAGGTAAGAGTTTG; K352E/H355P.sen: GTGACAGTCACGAACTCTTACGTGAATATGC; and K352E/H355P.antisen: GCATATTCAGGTAAGAGTTCGTGACTGTCAC. All mutants were confirmed by sequencing.

Statistics

We evaluated differences between groups with the Student's *t* test, and for multiple groups by two-way ANOVA, followed by *a post-hoc* Tukey's test. Significance was set at $p < 0.05$. All values are expressed as means ± SEM. All computations were performed using the Statgraphic statistical graphics system (STSC Inc., Rockville, MD, USA).

Abbreviations

GPRC6A: G protein-coupled receptor family C group 6 member A; Ocn: Osteocalcin; VFT: Venus Flytrap; CRD: Cysteine Rich Domain; TM: Transmembrane; cAMP: Cyclic adenosine monophosphate; ERK: extracellular-signal-regulated kinase.

Acknowledgements

Author Contributions

Rupesh Agarwal, Micholas Dean Smith and Jeremy C. Smith were responsible for the computational studies and Min Pi, Ruisong Ye, and L. Darryl Quarles were responsible for the experimental studies regarding GPRC6A. All authors contributed to the writing of the paper.

Funding and additional information

This work was supported by grants from the National Institutes of Health, NIAMS grant number 1R01AR071930 and NIDDK grant numbers 1R01DK121132 and 1R01DK120567 (LDQ)

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Figures and Tables

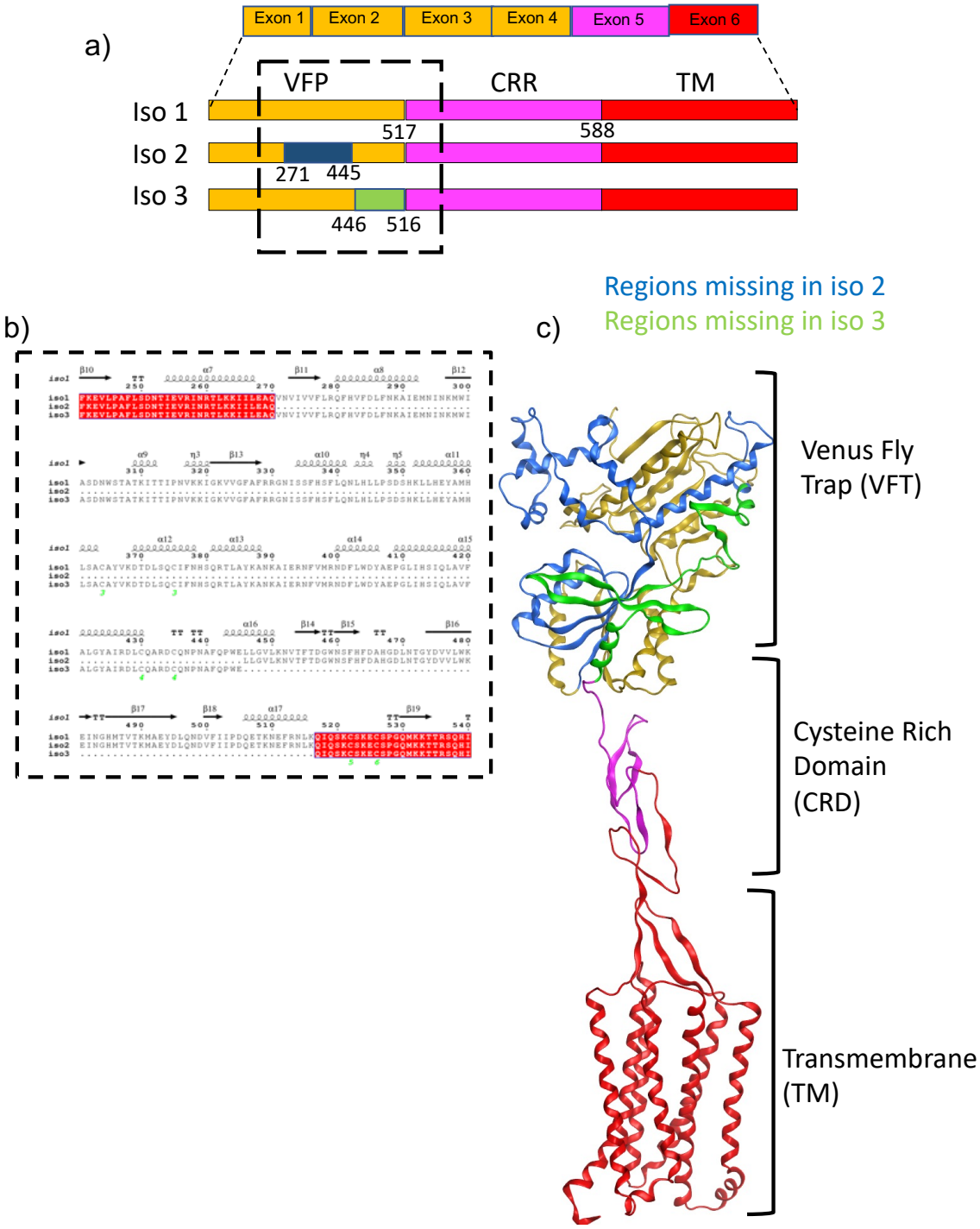


Figure 1. a) Placement of exons and its location in GPRC6A isoform 1, 2 and 3 structure; b) Sequence alignment of Venus flytrap (VFT) domain showing missing regions of isoform 2 and 3;

Structure model of GPRC6A showing the three domains and the missing region of isoform 2 and 3.

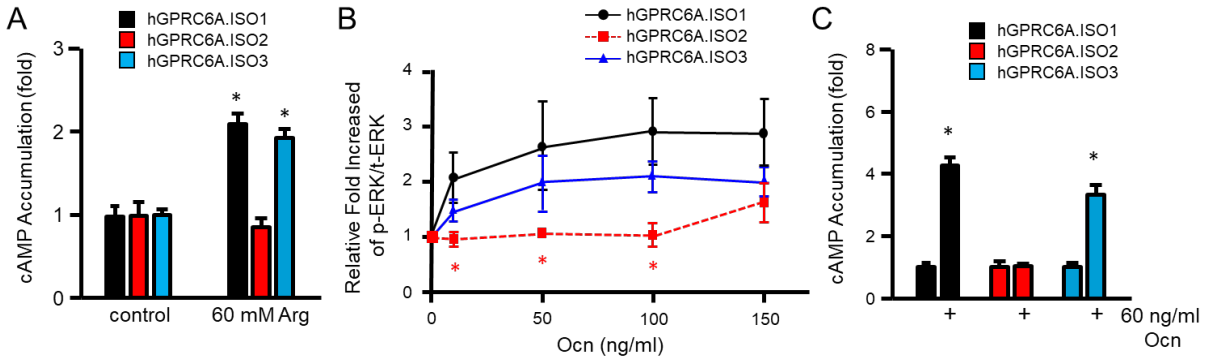


Figure 2. Activities of GPRC6A isoforms on response to Ocn. A. Comparison of effects of L-Arg on GPRC6A isoforms mediated cAMP accumulation. Comparison of dose-dependent effects of Ocn on GPRC6A isoforms mediated ERK activation (B) and cAMP accumulation (C). HEK-293 cells were transfected with cDNA plasmids of GPRC6A isoform 1, 2 or 3 for 48 hours, after incubated in Dulbecco's modified Eagle's medium /F-12 containing 0.1% bovine serum albumin quiescence media for 4 hours, then exposed to L-Arg or Ocn at indicated concentrations for 15 minutes for ERK activation, or 40 minutes for cAMP accumulation details as described under "Methods". * indicate a significant difference from control and stimulation groups at $p<0.05$.

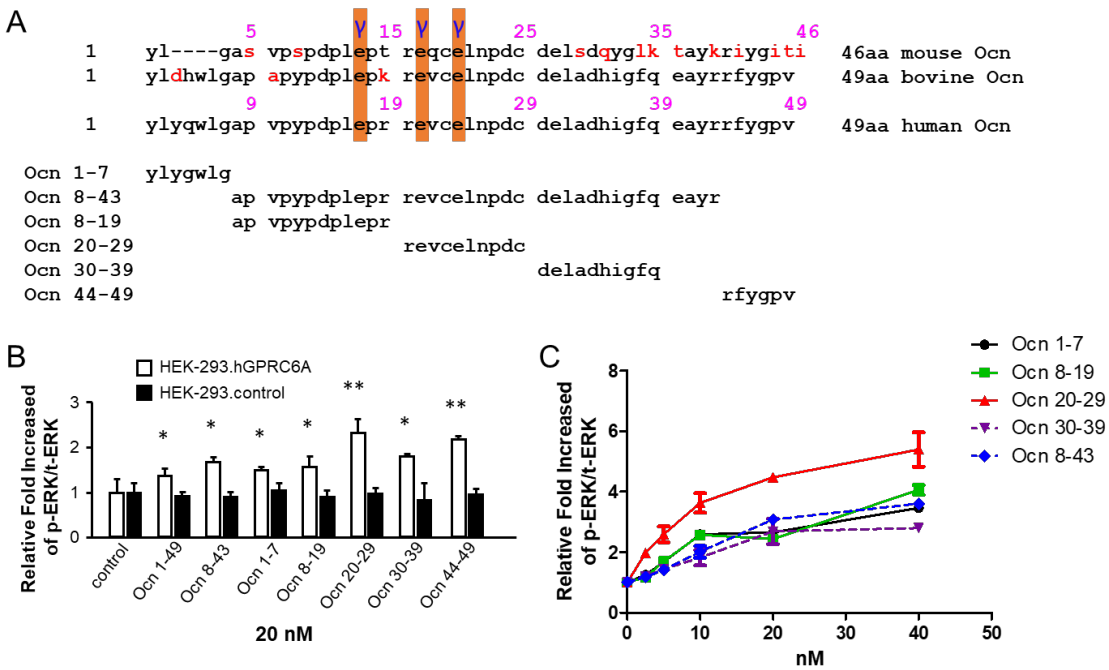


Figure 3. The activities of human Ocn peptide fragments. A, The sequence alignment of full length mouse, bovine and human Ocn (upper panel), and synthesized peptide fragments of human Ocn (bottom panel). Red alphabets indicate different amino acids compared to human Ocn sequence. "γ" indicates carboxylation sites in Ocn. B, Comparison of the activities of synthesized fragments of human Ocn (at 20 nM) by ERK phosphorylation in HEK-293 transfected with vector pcDNA3 (black bar) and HEK-293 cells transfected with pcDNA3-hGPRC6A cDNA (white bar). * and ** indicate a significant difference from control and stimulation groups at $p<0.05$ and $p<0.01$, respectively. C, The dose response of synthesized fragments of Ocn in HEK-293 cells transfected with pcDNA3-hGPRC6A cDNA. The ERK phosphorylation was measured 15 minutes for Ocn fragments at concentration as indicated at indicated stimulation in HEK-293 cells transfected with pcDNA3-hGPRC6A cDNA after 4 hours quiescence.

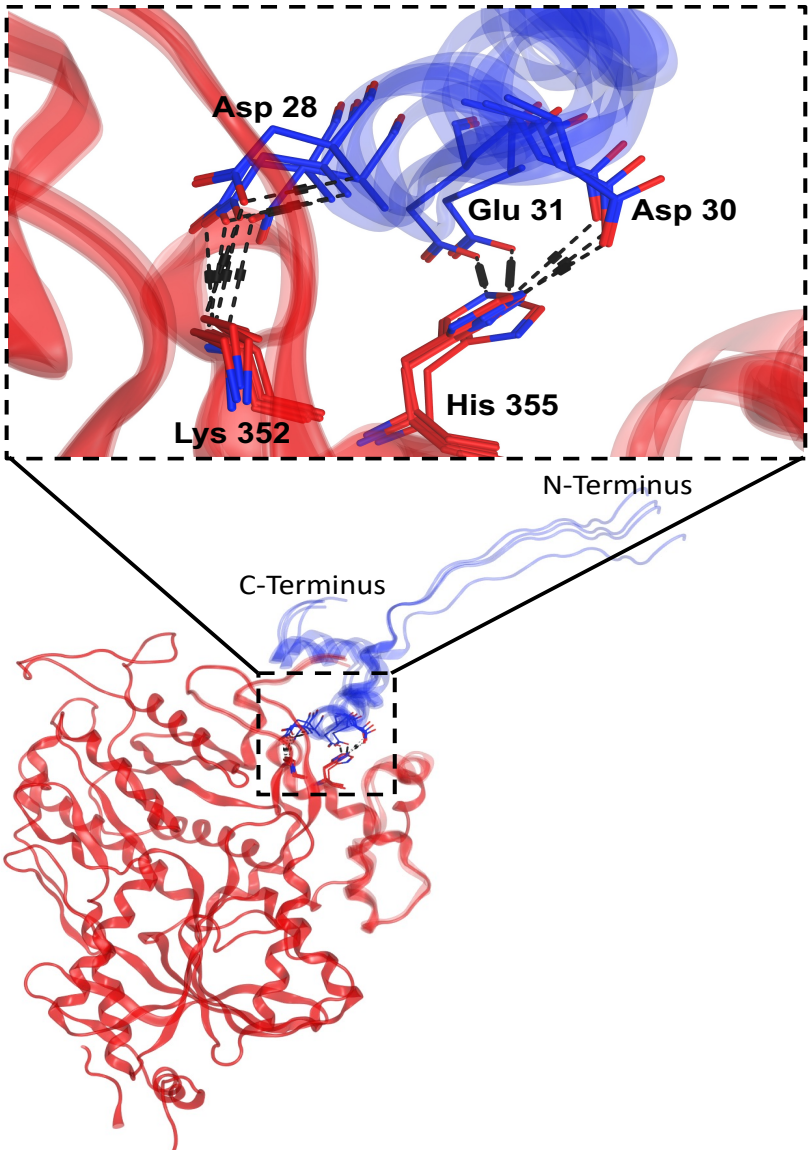


Figure 4: Superimposition of the top 4 models of GPRC6A's VFT domain (in red) and Ocn (in blue) complex generated from HADDOCK. The key residue interactions present in the common in helix region (352-363) of GPRC6A's VFT and residues 20-29 of Ocn.

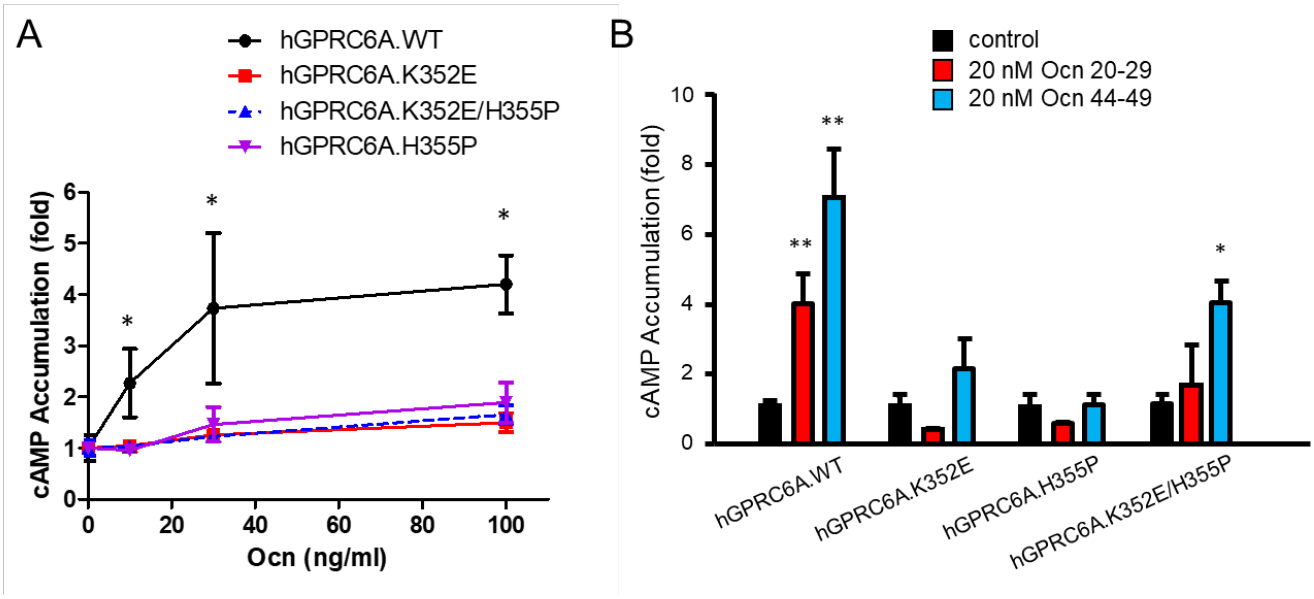


Figure 5. Mutagenesis of residues in predicted Ocn-binding pocket of GPRC6A. A, Comparison of dose-dependent effects of Ocn on cAMP accumulation in HEK-293 cells transfected with the cDNA plasmids of WT, K352E, H355P, or K352E/H355P hGPRC6A. * indicate a significant difference from control and stimulation groups at $p<0.05$ and $p<0.01$, respectively. B, Comparison of Ocn small fragments, Ocn 20-29 or Ocn 44-49 at 20 nM concentration on cAMP accumulation in HEK-293 cells transfected with the cDNA plasmids of WT, K352E, H355P, or K352E/H355P hGPRC6A. * and ** indicate a significant difference from control and stimulation groups at $p<0.05$ and $p<0.01$, respectively.

414 **Table 1: List of hydrogen bonds between GPRC6A-VFT and Ocn in best 4 structures from**
415 **the top cluster from HADDOCK.**
416

GPRC6A residue id	Ocn residue id	Distance (Å)	Model number
Asn389	Asp34	2.79	1
Asp26	Tyr46	2.73	1
His355	Asp30	2.93	1
Lys352	Asp28	3.31	1
Ser348	Glu31	2.67	1
Asp26	Asn26	2.81	2
Asp26	Phe45	3.83	2
Asp26	Tyr46	2.73	2
Gln381	Glu24	2.8	2
Gln381	Pro27	3.06	2
His355	Asp30	3.38	2
His355	Glu31	2.8	2
Lys352	Asp28	3.12	2
Ser348	Glu31	2.71	2
Ala385	Asp30	3.47	3
Asn389	Asp34	2.81	3
Asn389	His35	4.7	3
Asp26	Asn26	2.81	3
Gln19	Pro18	3.89	3
Gln19	Glu21	3.03	3
Gln381	Glu24	2.81	3
His355	Asp30	2.91	3
Lys352	Asp28	3.18	3
Pro20	Pro18	3.48	3
Ser348	Glu31	3.6	3
Ala385	Asp30	3.57	4
Asp26	Phe45	3.38	4
Gln19	Pro18	3.86	4
Gln381	Glu24	2.8	4
Gln381	Asp30	2.87	4
His355	Glu31	2.75	4
Lys352	Asp28	3.06	4
Pro20	Pro18	3.06	4
Ser348	Glu31	2.71	4

417

Supporting Information

Table S1. The amino acid sequence of synthesized human Ocn.

Name	Sequence
Ocn 1-7	YLYGWLG
Ocn 8-43	APVPYPDPLEPRREVCELNPDCDELADHIGFQEAYR
Ocn 8-19	APVPYPDPLEPR
Ocn 20-29	REVCELNPDC
Ocn 30-39	DELADHIGFQ
Ocn 44-49	RFYGPV

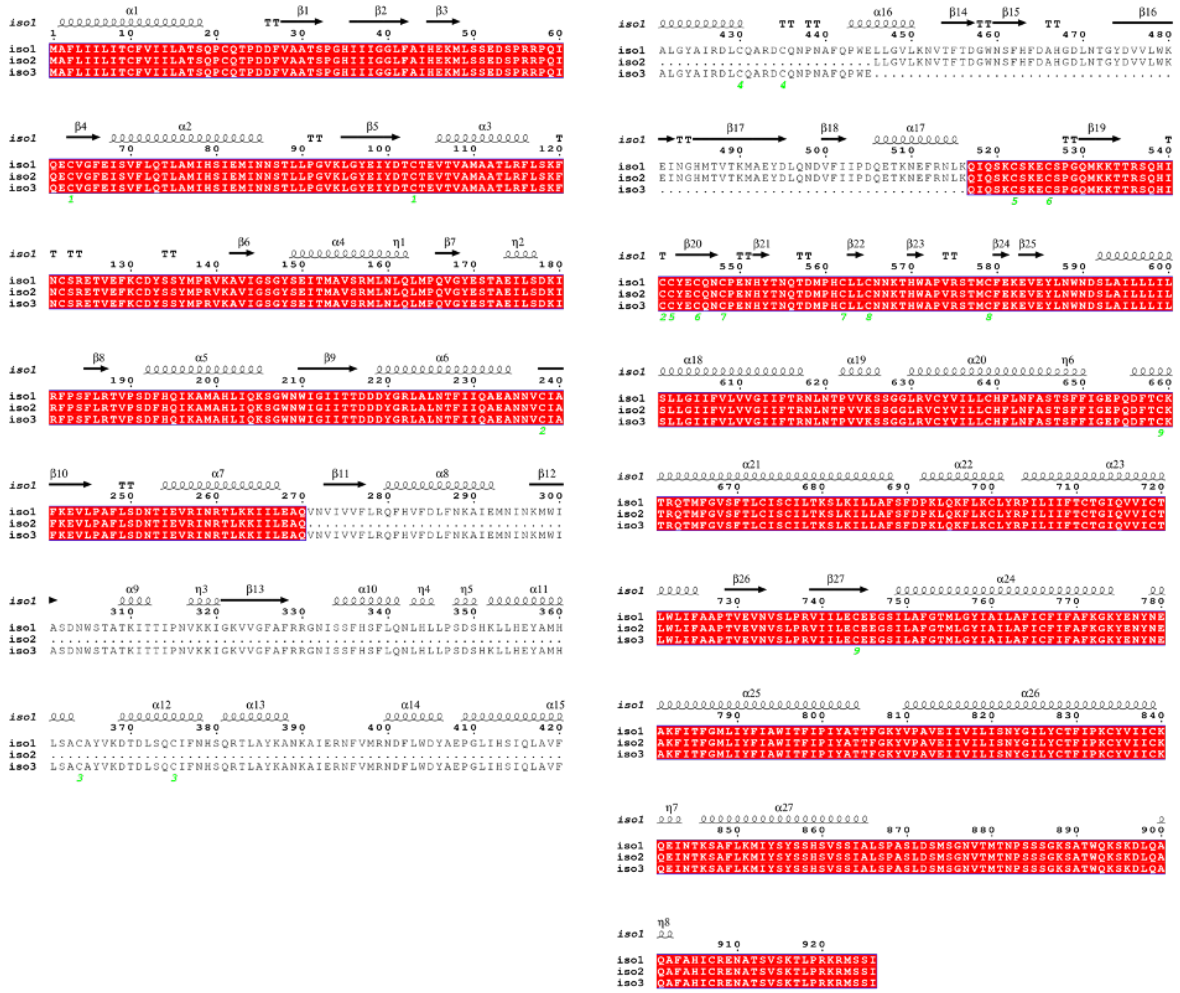
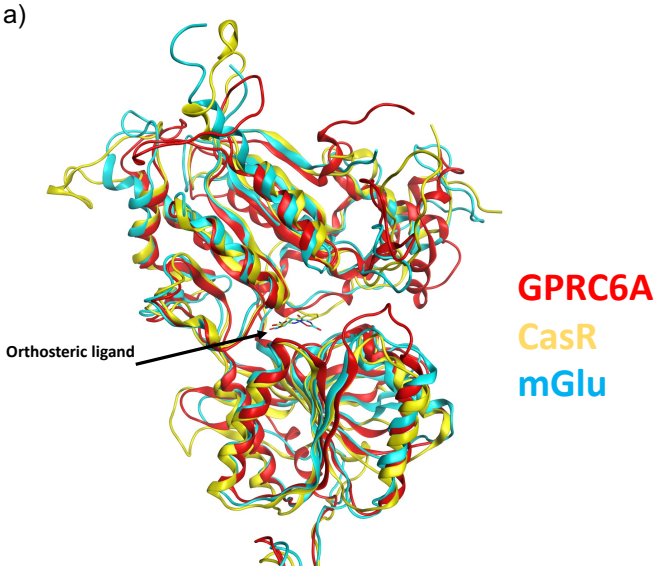


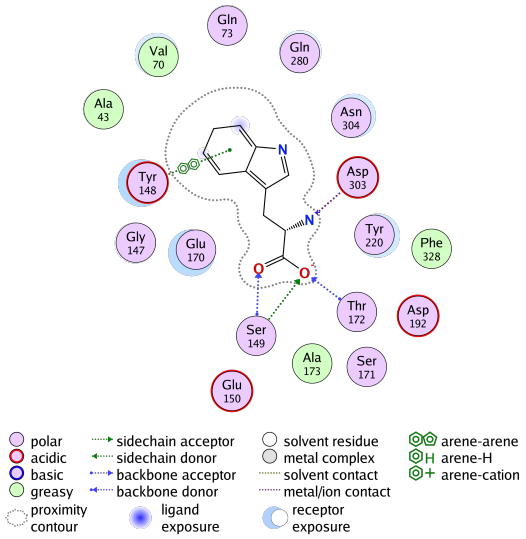
Figure S1: Sequence alignment of GPRC6A isoform 1, 2 and 3.

a)

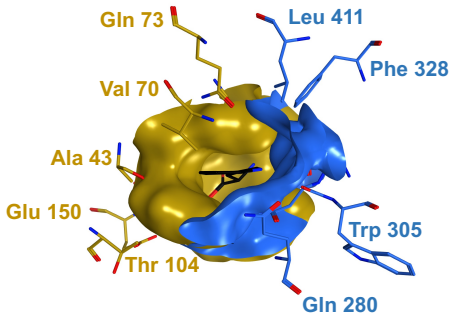


434

b)



c)

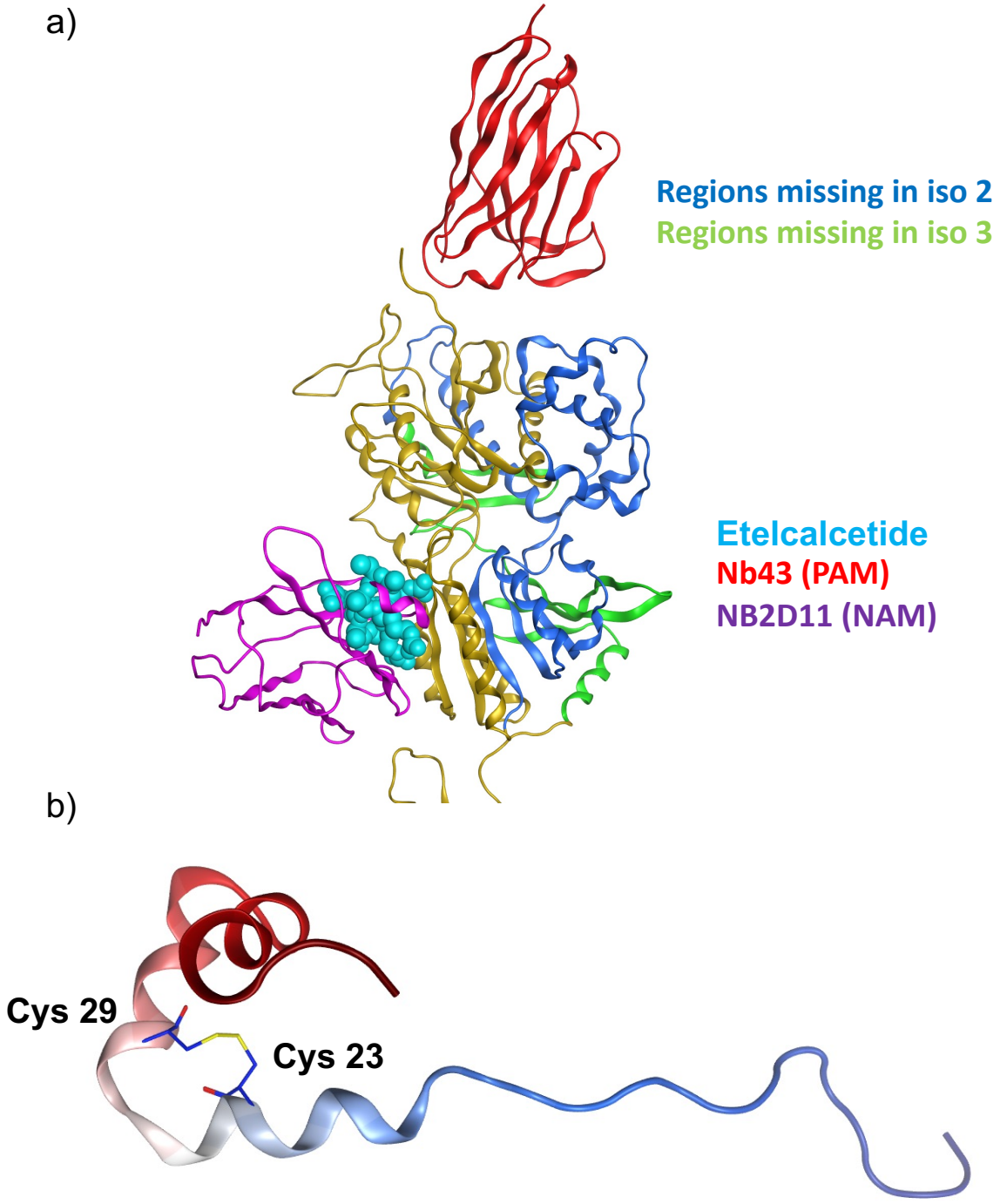


Residues missing in iso 2: 271-445

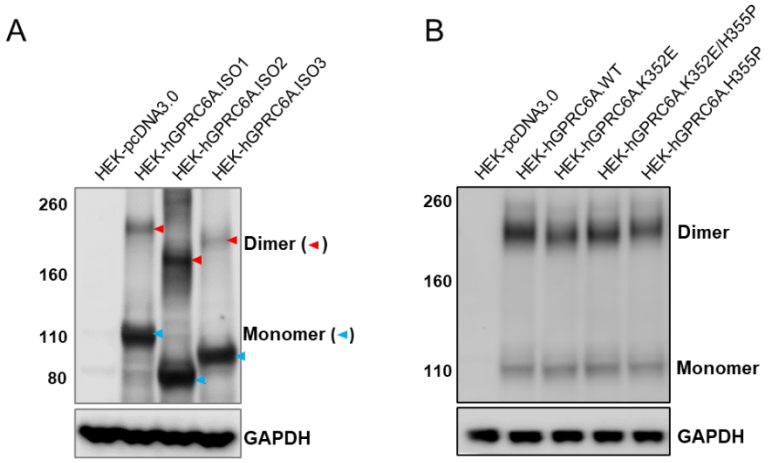
435

436

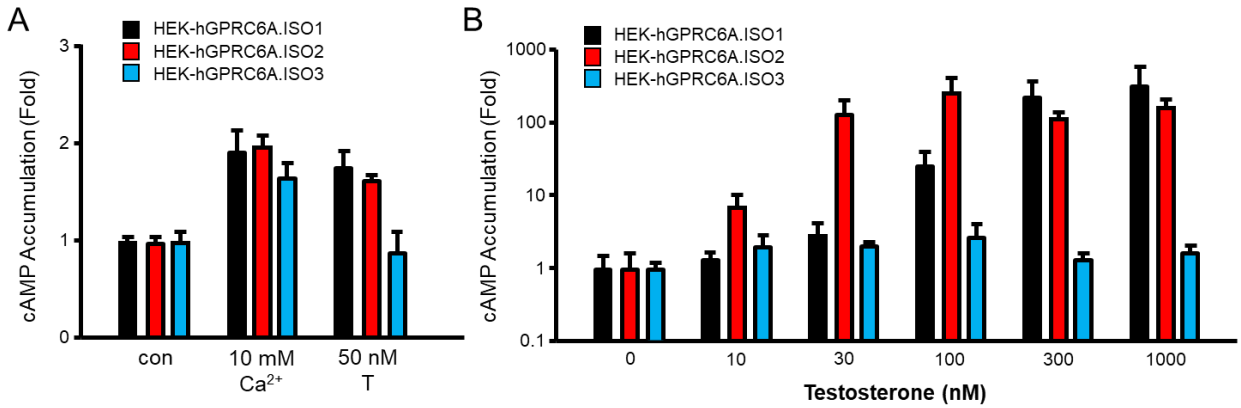
Figure S2: a) Superimposed structure of AlphaFold2 minimized model of GPRC6A, CasR with L-Trp (PDB 7DTU) and mGlu with orthosteric analogue L-quisqualate (PDB 6N51) (in ribbon) showing orthosteric ligands. b) Interaction map of L-Trp in isoform 1 c) Modeling of the orthosteric ligand from CasR in GPRC6A-isoform 1. Black: Trp (placed using CasR:Trp as template); Blue (orthosteric residues/ orthosteric binding pocket region missing in iso2); yellow: iso1 orthosteric pocket region.



445 Figure S3: a) Superimposition of crystal structures on GPRC6A model. Yellow ribbon: GPRC6A;
446 Dark blue ribbon: region missing in iso2; green ribbon: part missing in iso3; Red ribbon: PAM
447 Nb43 from mGlu (PDB 6N4Y); magenta ribbon: NAM NB2D11 from CaSR (PDB 7E6U) ; Cyan
448 sphere: etelcalcetide (calcimimetic drug) (PDB 7M3G); b) AlphaFold2 minimized model of hOcn
449 colored from N to C terminus (Blue to red) showing disulfide bond between Cys23 and Cys 29
450
451
452



453
454 Figure S4. The protein expression of GPRC6A isoforms (A) and mutants (B**).**
455
456
457



C

```
hGPRC6A     VAMAATLRFLSKFNCSR-ETVEFKCDYSSYMPRVKAVIGSGYSEITMAVSRMLNLQ
mGPRC6A     AAMAATLRFLSKFNCSR-ETVVFQCDYSSYMPRVKAVIGAGYSETSIAVSRMLNLQ
hCASR       KALEATLSFVAQNKIDSLNLDEFCNC-SEHIPSTIAVVGATGSGVSTAVANLLGLF
```

Exon 2 | Exon 3

```
hGPRC6A     LMPQVGYESTAEILSDKIRFPSFLRTVPSDFHQIKAMAHLIQKSGWNWIGIITTD
mGPRC6A     LMPQVSYESTAEILSDKIRFPSFLRTVPSDFYQTKAMAHLIRQSGWNWIGAITTD
hCASR       YIPQVSYASSSRLLSNKNQFKSFLRTIPNDEHQATAMADIIEYFRWNWVGTIAAD
```

Figure S5. GPRC6A isoform 3 loss responsive to T activation. A. Comparison of effects of Ca^{2+} or T on GPRC6A isoforms mediated cAMP accumulation. B. Comparison of dose-dependent effects of T on GPRC6A isoforms mediated cAMP accumulation. HEK-293 cells were transfected with cDNA plasmids of GPRC6A isoform 1, 2 or 3 for 48 hours, after incubated in Dulbecco's modified Eagle's medium /F-12 containing 0.1% bovine serum albumin quiescence media for 4 hours, then exposed to Ca^{2+} or T at indicated concentrations for 40 minutes for cAMP accumulation details as described under "Methods". * and ** indicate a significant difference from control and stimulation groups at $p<0.05$, and 0.01. C. Alignment of human GPRC6A, mouse GPRC6A and human CaSR showing pedieted Ca2+ binding sites. Red arrow shows the junction of exon 2 and exon 3. ♦ indicate expected amino acid in GPRC6A depended by CaSR.

References

1. Lee, N. K., Sowa, H., Hinoi, E., Ferron, M., Ahn, J. D., Confavreux, C., Dacquin, R., Mee, P. J., McKee, M. D., Jung, D. Y., Zhang, Z., Kim, J. K., Mauvais-Jarvis, F., Ducy, P., and Karsenty, G. (2007) Endocrine regulation of energy metabolism by the skeleton. *Cell* **130**, 456-469
2. Oury, F., Sumara, G., Sumara, O., Ferron, M., Chang, H., Smith, C. E., Hermo, L., Suarez, S., Roth, B. L., Ducy, P., and Karsenty, G. (2011) Endocrine regulation of male fertility by the skeleton. *Cell* **144**, 796-809

481 3. Pi, M., Nishimoto, S. K., and Darryl Quarles, L. (2021) Explaining Divergent Observations
482 Regarding Osteocalcin/GPRC6A Endocrine Signaling. *Endocrinology* **162**
483 4. Pi, M., Nishimoto, S. K., and Quarles, L. D. (2017) GPRC6A: Jack of all metabolism (or
484 master of none). *Mol Metab* **6**, 185-193
485 5. Pi, M., Wu, Y., and Quarles, L. D. (2011) GPRC6A mediates responses to osteocalcin in
486 beta-cells in vitro and pancreas in vivo. *J Bone Miner Res* **26**, 1680-1683
487 6. Pi, M., Xu, F., Ye, R., Nishimoto, S. K., Williams, R. W., Lu, L., and Darryl Quarles, L. (2020)
488 Role of GPRC6A in Regulating Hepatic Energy Metabolism in Mice. *Sci Rep* **10**, 7216
489 7. Mera, P., Laue, K., Wei, J., Berger, J. M., and Karsenty, G. (2016) Osteocalcin is necessary
490 and sufficient to maintain muscle mass in older mice. *Mol Metab* **5**, 1042-1047
491 8. Mukai, S., Mizokami, A., Otani, T., Sano, T., Matsuda, M., Chishaki, S., Gao, J., Kawakubo-
492 Yasukochi, T., Tang, R., Kanematsu, T., Takeuchi, H., Jimi, E., and Hirata, M. (2021)
493 Adipocyte-specific GPRC6A ablation promotes diet-induced obesity by inhibiting
494 lipolysis. *J Biol Chem* **296**, 100274
495 9. Mizokami, A., Mukai, S., Gao, J., Kawakubo-Yasukochi, T., Otani, T., Takeuchi, H., Jimi, E.,
496 and Hirata, M. (2020) GLP-1 signaling is required for improvement of glucose tolerance
497 by osteocalcin. *J Endocrinol* **244**, 285-296
498 10. Mizokami, A., Yasutake, Y., Higashi, S., Kawakubo-Yasukochi, T., Chishaki, S., Takahashi,
499 I., Takeuchi, H., and Hirata, M. (2014) Oral administration of osteocalcin improves
500 glucose utilization by stimulating glucagon-like peptide-1 secretion. *Bone* **69**, 68-79
501 11. Pi, M., Kapoor, K., Ye, R., Nishimoto, S. K., Smith, J. C., Baudry, J., and Quarles, L. D.
502 (2016) Evidence for Osteocalcin Binding and Activation of GPRC6A in beta-Cells.
503 *Endocrinology* **157**, 1866-1880
504 12. Pi, M., Kapoor, K., Ye, R., Hwang, D. J., Miller, D. D., Smith, J. C., Baudry, J., and Quarles,
505 L. D. (2018) Computationally identified novel agonists for GPRC6A. *PLoS One* **13**,
506 e0195980
507 13. Pi, M., Kapoor, K., Wu, Y., Ye, R., Senogles, S. E., Nishimoto, S. K., Hwang, D. J., Miller, D.
508 D., Narayanan, R., Smith, J. C., Baudry, J., and Quarles, L. D. (2015) Structural and
509 Functional Evidence for Testosterone Activation of GPRC6A in Peripheral Tissues. *Mol*
510 *Endocrinol* **29**, 1759-1773
511 14. Pi, M., Kapoor, K., Ye, R., Smith, J. C., Baudry, J., and Quarles, L. D. (2018) GPCR6A Is a
512 Molecular Target for the Natural Products Gallate and EGCG in Green Tea. *Mol Nutr*
513 *Food Res* **62**, e1700770
514 15. Pin, J. P., Galvez, T., and Prezeau, L. (2003) Evolution, structure, and activation
515 mechanism of family 3/C G-protein-coupled receptors. *Pharmacol Ther* **98**, 325-354
516 16. Koehl, A., Hu, H., Feng, D., Sun, B., Zhang, Y., Robertson, M. J., Chu, M., Kobilka, T. S.,
517 Laeremans, T., Steyaert, J., Tarrasch, J., Dutta, S., Fonseca, R., Weis, W. I., Mathiesen, J.
518 M., Skiniotis, G., and Kobilka, B. K. (2019) Structural insights into the activation of
519 metabotropic glutamate receptors. *Nature* **566**, 79-84
520 17. Chen, X., Wang, L., Cui, Q., Ding, Z., Han, L., Kou, Y., Zhang, W., Wang, H., Jia, X., Dai, M.,
521 Shi, Z., Li, Y., Li, X., and Geng, Y. (2021) Structural insights into the activation of human
522 calcium-sensing receptor. *Elife* **10**
523 18. Park, J., Zuo, H., Frangaj, A., Fu, Z., Yen, L. Y., Zhang, Z., Mosyak, L., Slavkovich, V. N., Liu,
524 J., Ray, K. M., Cao, B., Vallese, F., Geng, Y., Chen, S., Grassucci, R., Dandey, V. P., Tan, Y.

525 Z., Eng, E., Lee, Y., Kloss, B., Liu, Z., Hendrickson, W. A., Potter, C. S., Carragher, B.,
526 Graziano, J., Conigrave, A. D., Frank, J., Clarke, O. B., and Fan, Q. R. (2021) Symmetric
527 activation and modulation of the human calcium-sensing receptor. *Proc Natl Acad Sci U*
528 *S A* **118**
529 19. Wen, T., Wang, Z., Chen, X., Ren, Y., Lu, X., Xing, Y., Lu, J., Chang, S., Zhang, X., Shen, Y.,
530 and Yang, X. (2021) Structural basis for activation and allosteric modulation of full-
531 length calcium-sensing receptor. *Sci Adv* **7**
532 20. Gao, Y., Robertson, M. J., Rahman, S. N., Seven, A. B., Zhang, C., Meyerowitz, J. G.,
533 Panova, O., Hannan, F. M., Thakker, R. V., Brauner-Osborne, H., Mathiesen, J. M., and
534 Skiniotis, G. (2021) Asymmetric activation of the calcium-sensing receptor homodimer.
535 *Nature* **595**, 455-459
536 21. Wellendorph, P., and Brauner-Osborne, H. (2004) Molecular cloning, expression, and
537 sequence analysis of GPRC6A, a novel family C G-protein-coupled receptor. *Gene* **335**,
538 37-46
539 22. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O.,
540 Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.
541 A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T.,
542 Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M.,
543 Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K.,
544 Kohli, P., and Hassabis, D. (2021) Highly accurate protein structure prediction with
545 AlphaFold2. *Nature* **596**, 583-589
546 23. Jumper, J., and Hassabis, D. (2022) Protein structure predictions to atomic accuracy with
547 AlphaFold2. *Nat Methods* **19**, 11-12
548 24. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D.,
549 Stroe, O., Wood, G., Laydon, A., Zidek, A., Green, T., Tunyasuvunakool, K., Petersen, S.,
550 Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N.,
551 Kohli, P., Kleywegt, G., Birney, E., Hassabis, D., and Velankar, S. (2022) AlphaFold2
552 Protein Structure Database: massively expanding the structural coverage of protein-
553 sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444
554 25. Skolnick, J., Gao, M., Zhou, H., and Singh, S. (2021) AlphaFold2 2: Why It Works and Its
555 Implications for Understanding the Relationships of Protein Sequence, Structure, and
556 Function. *J Chem Inf Model* **61**, 4827-4831
557 26. Baumgrass, R., Williamson, M. K., and Price, P. A. (1997) Identification of peptide
558 fragments generated by digestion of bovine and human osteocalcin with the lysosomal
559 proteinases cathepsin B, D, L, H, and S. *J Bone Miner Res* **12**, 447-455
560 27. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein
561 docking approach based on biochemical or biophysical information. *J Am Chem Soc* **125**,
562 1731-1737
563 28. Rueda, P., Harley, E., Lu, Y., Stewart, G. D., Fabb, S., Diepenhorst, N., Cremers, B.,
564 Rouillon, M. H., Wehrle, I., Geant, A., Lamarche, G., Leach, K., Charman, W. N.,
565 Christopoulos, A., Summers, R. J., Sexton, P. M., and Langmead, C. J. (2016) Murine
566 GPRC6A Mediates Cellular Responses to L-Amino Acids, but Not Osteocalcin Variants.
567 *PLoS One* **11**, e0146846

29. Power, M. J., and Fottrell, P. F. (1991) Osteocalcin: diagnostic methods and clinical applications. *Crit Rev Clin Lab Sci* **28**, 287-335
30. Teng, B., Huang, C., Cheng, C. L., Udduttula, A., Yu, X. F., Liu, C., Li, J., Yao, Z. Y., Long, J., Miao, L. F., Zou, C., Chu, J., Zhang, J. V., and Ren, P. G. (2020) Newly identified peptide hormone inhibits intestinal fat absorption and improves NAFLD through its receptor GPRC6A. *J Hepatol* **73**, 383-393
31. Jeong, E., Kim, Y., Jeong, J., and Cho, Y. (2021) Structure of the class C orphan GPCR GPR158 in complex with RGS7-Gbeta5. *Nat Commun* **12**, 6805
32. Varadi, M., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Anyango, S., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Deshpande, M., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Nair, S., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Natassia, C., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Yordanova, G., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Yuan, D., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Stroe, O., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Wood, G., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Laydon, A., DeepMind, L., UK, Žídek, A., DeepMind, L., UK, Green, T., DeepMind, L., UK, Tunyasuvunakool, K., DeepMind, L., UK, Petersen, S., DeepMind, L., UK, Jumper, J., DeepMind, L., UK, Clancy, E., DeepMind, L., UK, Green, R., DeepMind, L., UK, Vora, A., DeepMind, L., UK, Lutfi, M., DeepMind, L., UK, Figurnov, M., DeepMind, L., UK, Cowie, A., DeepMind, L., UK, Hobbs, N., DeepMind, L., UK, Kohli, P., DeepMind, L., UK, Kleywegt, G., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Birney, E., European Molecular Biology Laboratory, E. B. I., Hinxton, UK, Hassabis, D., DeepMind, L., UK, Velankar, S., and European Molecular Biology Laboratory, E. B. I., Hinxton, UK. (2022) AlphaFold2 Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* **50**
33. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *Journal of computational chemistry* **26**, 1668-1688
34. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696-3713
35. Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **98**, 10089-10092
36. Price, D. J., and Brooks, C. L., 3rd. (2004) A modified TIP3P water potential for simulation with Ewald summation. *J Chem Phys* **121**, 10096-10103
37. Miyamoto, S., and Kollman, P. A. (1992) Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *Journal of computational chemistry* **13**, 952-962
38. Zwanzig, R. (1973) Nonlinear generalized Langevin equations. *Journal of Statistical Physics* **9**, 215-220
39. Åqvist, J., Wennerström, P., Nervall, M., Bjelic, S., and Brandsdal, B. O. (2004) Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chemical physics letters* **384**, 288-294

612 40. Nishimoto, S. K., and Price, P. A. (1979) Proof that the gamma-carboxyglutamic acid-
613 containing bone protein is synthesized in calf bone. Comparative synthesis rate and
614 effect of coumadin on synthesis. *J Biol Chem* **254**, 437-441
615 41. Sabek, O. M., Nishimoto, S. K., Fraga, D., Tejpal, N., Ricordi, C., and Gaber, A. O. (2015)
616 Osteocalcin Effect on Human beta-Cells Mass and Function. *Endocrinology* **156**, 3137-
617 3146
618 42. Nishimoto, S. K. (1990) A colorimetric assay specific for gamma-carboxyglutamic acid-
619 containing proteins: its utility in protein purification procedures. *Anal Biochem* **186**, 273-
620 279
621 43. Nishimoto, S. K., and Price, P. A. (1985) The vitamin K-dependent bone protein is
622 accumulated within cultured osteosarcoma cells in the presence of the vitamin K
623 antagonist warfarin. *J Biol Chem* **260**, 2832-2836
624 44. Pi, M., Parrill, A. L., and Quarles, L. D. (2010) GPRC6A mediates the non-genomic effects
625 of steroids. *J Biol Chem* **285**, 39953-39964
626