Structural insights into ligand recognition and activation of the melanocortin-4 receptor

Huibing Zhang[1,2,3,4,#], Li-Nan Chen[1,2,3,4,#], Dehua Yang[5,6,7,#], Chunyou Mao[1,2,3,4], Qingya Shen[1,2,3,4], Wenbo Feng[8], Dan-Dan Shen[1,2,3,4], Antao Dai[5,6], Shanshan Xie[9], Yan Zhou[5,6], Jiao Qin[1,2,3,4], Jinpeng Sun[10,11], Daniel H. Scharf[12], Tingjun Hou[13], Tianhua Zhou[9], Ming-Wei Wang[5,6,7,8,14,15,*] , Yan Zhang[1,2,3,4,*]

[1]Department of Biophysics and Department of Pathology of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310058, Zhejiang, China

[2]Liangzhu Laboratory, Zhejiang University Medical Center, Hangzhou 311121, China

[3]MOE Frontier Science Center for Brain Research and Brain-Machine Integration, Zhejiang University School of Medicine, Hangzhou 310058, Zhejiang, China

[4]Key Laboratory of Immunity and Inflammatory Diseases of Zhejiang Province, Hangzhou 310058, Zhejiang, China

[5]The CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

[6]The National Center for Drug Screening, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

[7]University of Chinese Academy of Sciences, Beijing 100049, China.

[8]School of Pharmacy, Fudan University, Shanghai, China

[9]Department of Cell Biology and Cancer Institute of the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China; Institute of Gastroenterology, Zhejiang University, Hangzhou, China; Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, Zhejiang, China; Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

[10]Department of Physiology and Pathophysiology, School of Basic Medical Sciences, Peking University; Key Laboratory of Molecular Cardiovascular Science, Ministry of Education, Beijing, China

[11]Key Laboratory Experimental Teratology of the Ministry of Education, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Cheeloo College of Medicine, Shandong University, Shandong, China

[12]Department of Microbiology and The Children's Hospital, Zhejiang University School of Medicine, Hangzhou 310058, Zhejiang, China

[13]Innovation Institute for Artificial Intelligence in Medicine of Zhejiang University, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, Zhejiang, China

[14]School of Life Science and Technology, ShanghaiTech University, Shanghai, China

[15]School of Basic Medical Sciences, Fudan University, Shanghai, China

[#]These authors contributed equally to this work.

[*]Correspondence: zhang_yan@zju.edu.cn (Y.Z.); mwwang@simm.ac.cn (M.-W.W.)

SUMMARY

Melanocortin-4 receptor (MC4R) plays a central role in the regulation of energy homeostasis. Its high sequence similarity to other MC receptor family members, low agonist selectivity and the lack of structural information concerning receptor activation have hampered the development of MC4R-seletive therapeutics to treat obesity. Here, we report four high-resolution structures of full-length MC4R in complex with the heterotrimeric G_s protein stimulated by the endogenous peptide α-MSH, FDA-approved drugs afamelanotide (Scenesse™) and bremelanotide (Vyleesi™), and selective small-molecule ligand THIQ, respectively. Together with pharmacological studies, our results reveal the conserved binding mode of peptidic agonists, the distinctive molecular details of small-molecule agonist recognition underlying receptor subtype selectivity, and distinct activation mechanism for MC4R, thereby offering new insights into G protein coupling. Our work may facilitate the discovery of selective therapeutic agents targeting MC4R.

INTRODUCTION

Melanocortin-4 receptor (MC4R), a key component of the leptin-melanocortin pathway, acts at the intersection of homeostatic maintenance of energetic state.[1-6] Mutations in the MC4R gene were identified among patients with severe obesity from early childhood.[7,8] By far, natural occurring loss of function (LOF) MC4R mutations are the most frequent monogenic cause of obesity or binge eating disorder.[9] MC4R agonists have been shown efficacious in appetite suppression, food intake reduction and body weight loss in people with high BMI (body mass index) caused by leptin, proopiomelanocortin (POMC) or MC4R deficiency.[10] Obviously, MC4R plays a central role in regulating energy balance and satiety, and is one of the best validated drug targets for obesity.[11] MC4R belongs to the melanocortin receptor family, a group of five (MC1R − MC5R) class A G protein-coupled receptors (GPCRs) and is primarily coupled to the stimulatory G protein (G_s) and increases intracellular cyclic adenosine monophosphate (cAMP) accumulation.[12] MC1R governs mammalian skin and hair color by regulating the production of melanin.[13] MC2R is mainly located in the adrenal

68 cortex and controls the release of glucocorticoids.[2] MC3R participates in the control of

69 energy homeostasis and is implicated in immune responses, natriuresis and circadian

70 rhythm.[14] MC5R is involved in exocrine gland dysfunction, sebum regulation, and the

71 pathogenesis of acne.[15-17] Except MC2R, which is only activated by ACTH

72 (adrenocorticotropic hormone) and requires an interplay with the melanocortin receptor

73 accessory protein (MRAP) to attain functionality,[18] the other four MCRs are

74 endogenously activated by α-, β-, and γ-melanocyte-stimulating hormones (MSH), and

75 blocked by the natural inverse agonist agouti-related peptide (AgRP).[19] These features

76 put forward a keen request for the development of a selective ligand for MC4R.

77 The endogenous ligands for MCRs contain a conserved His-Phe-Arg-Trp (HFRW)

78 motif necessary for receptor activation (Supplementary information, Table S1).[20,21]

79 Among them, α-MSH (Ac-Ser-Tyr-Ser-Met-Glu-His-Phe-Arg-Trp-Gly-Lys-Pro-Val-NH$_2$)

80 is the most well-studied peptide that has pleiotropic functions in pigmentation,

81 cardiovascular system, erectile ability, energy homeostasis and immune responses.[21-

82 24] Driven by the pervasive issue of human obesity, tremendous efforts have been made

83 to design potent and selective therapeutic agents targeting MC4R, both of peptidic and

84 small-molecule nature.[23-27] The linear peptide afamelanotide, Ac-Ser-Tyr-Ser-Nle^4-Glu-

85 His-D-Phe^7-Arg-Trp-Gly-Lys-Pro-Val-NH$_2$ (also known as [Nle^4, DPhe^7]α-MSH, NDP-

86 α-MSH, melanotan-I, and MT-I), is capable of enhancing metabolic stability with

87 markedly improved potency compared to α-MSH.[21,28] Subsequent efforts in obtaining

88 peptides with better bioactivities led to the discovery of a cyclic peptide agonist, Ac-

89 Nle^4-c [Asp^5, D-Phe^7, Lys^{10}] α-MSH$_{(4-10)}$-NH$_2$ (melanotan-II, MT-II).[21,22] Bremelanotide

90 (C-terminal OH-MT-II) was approved in 2019 by the FDA to treat premenopausal

91 women with sexual excitement dysfunction.[29] Afamelanotide binds to MC1R in dermal

92 cells to ramp up melanin production, effectively tanning the skin, and as a photo

93 protectant received market approval in the European Union in 2014 and by the FDA in

94 2019.[30]

95 Unfortunately, owning to the high sequence homology among MCRs, most

96 peptidic agonists initially developed for MC4R lack the capability of avoiding cross-

97 reactivity with other MCR subtypes, such as MC1R, leading to adverse effects, *e.g.,*

98   skin pigmentation and nausea.[10,23,29] Intriguingly, subtype selectivity of MC4R was first

99   achieved by a small-molecule agonist, THIQ, initially developed by Merck (1400-fold

100   *vs*. MC1R, 1200-fold *vs*. MC3R and 360-fold *vs*. MC5R),[25] that significantly decreases

101   food consumption and simulates erectile responses.

102   The activation mechanism of MC4R and how it recognizes peptidic and small-

103   molecule agonists are poorly understood, hampering the structure-guided drug

104   discovery. Here, we employed single-particle cryo-electron microscopy (cryo-EM) to

105   determine four structures of the human MC4R-G_s complexes: three bound to α-MSH,

106   afamelanotide or bremelanotide at the resolutions of 3.0 Å, 3.0 Å and 3.1 Å,

107   respectively; and one bound to THIQ at the resolution of 3.1 Å (Fig. 1; Supplementary

108   information, Fig. S1, S2, Table S2). These structures reveal the molecular basis of

109   ligand recognition and receptor activation for MC4R, shed light on the subtype

110   selectivity of MCRs, and provide templates for the rational design of novel therapeutics

111   targeting MC4R to treat obesity.

112

113   **RESULTS**

114   **Overall structures of agonist–MC4R–G_s complexes**

115   To obtain a homogenous and stable MC4R–G_s complex, full-length MC4R was cloned

116   into a pFastBac1 vector with a LgBiT and double maltose binding protein (MBP) affinity

117   tag at the C-terminus. A dominant negative Gα$_s$ (DNGα$_s$), Gβ1 and Gγ2 were co-

118   expressed with MC4R in insect cells stabilized by a NanoBiT tethering strategy.[31-33]

119   Subsequently, assembly of the MC4R–G_s complex on the membrane was stimulated

120   with an excess amount of agonist (α-MSH, bremelanotide, afamelanotide or THIQ (Fig.

121   1) and further stabilized by nanobody 35 (Nb35).[34] The complex was then purified

122   through affinity and size-exclusion chromatography to homogeneity, yielding a stable

123   complex suitable for single-particle cryo-EM analysis (Supplementary information, Fig.

124   S1a-b).

125   All four complexes, including α-MSH–MC4R–G_s–Nb35, afamelanotide–MC4R–

126   G_s–Nb35, bremelanotide–MC4R–G_s–Nb35, and THIQ–MC4R–G_s–Nb35, were imaged

127   under a Titan Krios microscope equipped with K2 summit direct detector, and the

128   structures were determined at the global resolutions of 3.0 Å, 3.0 Å, 3.1 Å, and 3.1 Å,

129   respectively (Supplementary information, Fig. S1c-f, S2a-d, Table S2). The high-

130   resolution density maps showed that the majority of side chains in the signaling

131   complexes were well resolved and allowed us to model most regions of MC4R from

132   residues G39^{N-term} to Y320$^{8.61}$ (superscript indicates the Ballesteros–Weinstein

133   numbering system for class A GPCR)[35] except G233^{ICL3} to R236^{ICL3} and D111^{ECL1} to

134   Q115^{ECL1}; α-MSH, afamelanotide, bremelanotide and THIQ, the G_s heterotrimer, and

135   Nb35 in the final models were refined against the corresponding density map with

136   excellent geometry (Fig. 1a-b; Supplementary information, Fig. S2). Additionally, a

137   divalent calcium ion (Ca^{2+}) was consistently observed in the peptide- and small-

138   molecule-bound structures (Supplementary information, Fig. S2e-f, right, green

139   sphere). Therefore, our structures could provide detailed information of the ligand

140   binding mode between peptidic/small-molecule ligands and the receptor helix bundle,

141   as well as the interface between G_s and MC4R, allowing us to examine conformational

142   changes of the receptor upon agonist binding and the receptor-transducer coupling.

143   The overall conformations of MC4R–Gs complexes are almost identical in all four

144   active structures (Cα root mean squared deviation (r.m.s.d.) < 0.3 Å), so we used the

145   well-resolved α-MSH–MC4R–G_s structure (Fig. 1a) for further analysis below unless

146   otherwise noted. Although MC4R adopts the canonical seven transmembrane (7TM)

147   spanning helices architecture, it also displays several distinctive structural features.

148   First, MC4R does not possess the conserved disulfide bond that usually connects the

149   extracellular tip of TM3 to ECL2 in most GPCRs.[36] However, the lack of a conserved

150   disulfide bond also occurs in other class A GPCRs, such as sphingosine-1-phosphate

151   receptors   S1P$_1$R-S1P$_5$R,   lysophosphatidic   acid   receptors   LPA$_1$R-LPA$_3$R   and

152   cannabinoid   receptors   CB$_1$R-CB$_2$R.[37-42] Second, three cysteine residues in ECL3

153   (C271, C277 and C279) are completely conserved among MCRs, two of which (C271

154   and C277) form an intra-ECL3 disulfide bond and C279 forms an additional disulfide

155   bond with C40 in the N terminus (Supplementary information, Fig. S2e) which was not

156   observed in the crystal structure of the antagonist SHU9119-bound thermostabilized

157   MC4R (tsMC4R) in the inactive state.[43] At the cytoplasmic side, the 7TM bundle of

158 MC4R created a cavity to accommodate the heterotrimeric G_s protein. Unique to cryo-

159 EM structures of MC4R–G_s complexes, several water molecules were trapped in the

160 receptor–G_s interface (Fig. 1a). Notably, one water molecule unprecedentedly bridges

161 R147$^{3.50}$ of the conserved DRY motif with E392$^{G.H5.24}$ (common Gα numbering system

162 [44]) in the C-terminal of Gα$_s$ (Fig. 1a).

163   Most class A aminergic, peptide and lipid GPCRs contain an elongated ECL2

164 segment consisting of over 15 amino acids and forming a β-sheet or helix lid, which

165 usually covers the orthosteric binding pocket (Supplementary information, Fig. S3a-i).

166 In contrast to these class A GPCRs, the ECL2 of MC4R is extremely short with only

167 three amino acids, resulting in a widely open and extensively solvent-accessible

168 orthosteric binding pocket.

169 **Recognition of α-MSH by MC4R**

170 The extensive interactions between the endogenous agonist α-MSH and MC4R

171 involve all seven TM helices plus the extended N terminus, ECL2 and ECL3 (Fig. 2a-

172 b; details are presented in Supplementary information, Table S3). Electrostatic surface

173 analysis revealed that the orthosteric binding pocket exhibits obvious amphiphilic

174 characteristics, with the TM 2-3-4 half negatively charged while the rest of the pocket

175 nearly hydrophobic (Supplementary information, Fig. S4a). The α-MSH inserts into the

176 binding pocket in a charged complementary fashion (Supplementary information, Fig.

177 S4a-c). The high-resolution density map showed an unambiguous density

178 corresponding to a Ca^{2+} ion, coordinated by two backbone carbonyl oxygen atoms of

179 Glu5 and Phe7 in α-MSH (peptide residues are labeled as three letters) and three

180 negatively charged residues (E100$^{2.60}$, D122$^{3.25}$ and D126$^{3.29}$) in MC4R (Fig. 2a). The

181 side chain of Tyr2 and Arg8 located in close proximity to D122$^{3.25}$, D126$^{3.29}$ and I185$^{4.61}$,

182 and Arg8 hydrogen bonded to S188 in ECL2 (Fig. 2a-b; details are presented in

183 Supplementary information, Table S3). In addition, F51$^{1.39}$, T101$^{2.61}$, I104$^{2.64}$, F284$^{7.35}$,

184 N285$^{7.36}$ and L288$^{7.39}$ prop up one supporting point His6, whereas I185$^{4.61}$, I194$^{5.40}$,

185 L197$^{5.43}$ and L265$^{6.55}$ accommodate another supporting point Trp9 of the U-shaped α-

186 MSH (Fig. 2a-b; details are presented in Supplementary information, Table S3). We

187 mutated these residues that contact α-MSH to alanine and observed a significant loss

188   of agonist-induced signal transduction in mutants F51$^{1.39}$A, E100$^{2.60}$A, D122$^{3.25}$A,

189   D126$^{3.29}$A, I185$^{4.61}$A, H264$^{6.54}$A and L288$^{7.39}$A, confirming the important roles of these

190   residues in α-MSH–induced receptor activation (Fig. 2c; details are presented in

191   Supplementary information, Table S3).

192   The EM density map indicated that Glu5 of α-MSH possibly adopted two distinct

193   rotamers, inducing local changes of the α-MSH–MC4R interface environment

194   (Supplementary information, Fig. S4g-h). One rotamer of Glu5 side chain is likely to

195   direct towards Lys11 forming a polar interaction, linking two flanks of the peptide. The

196   resulting cyclized ring-like conformation resembles the conformation of cyclic peptide

197   agonist bremelanotide (Fig. 3a, b). The other rotamer of Glu5 points down towards

198   His6. By adopting a similar direction, the aspartate at the corresponding site (Asp3) of

199   SHU9119 formed a hydrogen bond with His6 in the crystal structure of tsMC4R.[31] In

200   this case, His6 of α-MSH hydrogen bonded to N285$^{7.36}$ and T101$^{2.61}$ (Supplementary

201   information, Fig. S4g-h). Notably, tsMC4R contained a substitution of N97$^{2.57}$L, which

202   folded towards the bottom of the ligand biding pocket. Interestingly, we observed the

203   side chain of N97$^{2.57}$ in the α-MSH–MC4R structure adopting a conformation similar to

204   that in the tsMC4R structure (Supplementary information, Fig. S4g-h). In agreement

205   with the critical role of N97$^{2.57}$ in receptor activation, substitution of N97$^{2.57}$ with leucine

206   or alanine resulted in partial (about 25%) or complete loss of α-MSH–induced cAMP

207   responses (Supplementary information, Fig. S4i; details are presented in

208   Supplementary information, Table S4). N97$^{2.57}$A mutation resulted in near-complete

209   loss of surface trafficking (Supplementary information, Table S4), and consequent

210   dramatic reduction in activity (Supplementary information, Table S5) and a

211   corresponding cAMP response (Supplementary information, Fig. S4i). However,

212   although the mutant N97$^{2.57}$L maintained the robust surface expression level

213   (Supplementary information, Table S4) and binding affinity of α-MSH (Supplementary

214   information, Table S5), it partially lost cAMP response (Supplementary information, Fig.

215   S4i).

216

217   **Recognition of afamelanotide and bremelanotide by MC4R**

218 Sole replacement of Phe7 in α-MSH with D-Phe7 produced afamelanotide with

219 significant increase in binding affinity and functional activity.[21] The recognition mode of

220 afamelanotide by MC4R therefore is very similar to that of α-MSH (Fig. 3a, f;

221 Supplementary information, Fig. S5a-c). Bremelanotide has less interactions with

222 MC4R only involving TMs 2-7, owing to the omission of flanking residues at C and N

223 termini (Fig. 3b, c; Supplementary information, Fig. S5d; details are shown in

224 Supplementary information, Table S6, 7). Structural superimposition revealed that,

225 although substitution of Phe7 with D-Phe7 did not alter the overall shape of

226 afamelanotide compared to α-MSH, it indeed induced notable conformational changes

227 in the peptide. The N-terminal flank of afamelanotide shifted 1.5 Å away from TMs 2-3

228 measured at the Cα of Ser1, resulting in the partial loss of interactions with TM2 and

229 TM3 (Fig. 3a, f; Supplementary information, Fig. S5b, c). D-Phe7 almost stayed at the

230 same position as Phe7 of α-MSH. However, the replacement led to the neighboring

231 residue His6 swinging upward 1.4 Å to form a hydrogen bond with Glu5 (Fig. 3a),

232 together with the breaking of the hydrogen bond with $N97^{2.57}$ and π–π stacking with

233 $F51^{1.39}$ (Fig. 2a Supplementary information, Fig. S5b, c). Similar conformational

234 changes of Glu5 and His6 were also observed in the structure of bremelanotide-bound

235 MC4R (Fig. 3b, c, 3f; Supplementary information, Fig. S5d). Consistent with our

236 observation that $F51^{1.39}$ was not involved in the recognition of afamelanotide or

237 bremelanotide, the cAMP assay showed that the substitution of $F51^{1.39}$A only reduced

238 afamelanotide- and bremelanotide-induced responses by between 3- and 10-fold in

239 contrast to over 100-fold for α-MSH (Fig. 3d, e, details are shown in Supplementary

240 information, Table S8, 9). Additionally, the distance between Glu5 and His6 in

241 afamelanotide (3.4 Å) was in the preferable range of hydrogen bond compared to 4.6

242 Å in α-MSH (Fig. 3a). The more stabilized interaction between Glu5 and His6 in

243 afamelanotide trapping the side chain of Glu5 points toward His6 , rigidifying the local

244 conformation.

245

246 **Specific Recognition of THIQ by MC4R**

247 The cryo-EM map enabled the unambiguous assignment of tri-branched THIQ sitting

248   in the bottom of the orthosteric binding pocket (Fig. 1b, d; Supplementary information,

249   Fig. S5a). The binding of THIQ is mediated by residues from six transmembrane

250   helices (TMs 2-7) as well as residues from the N terminus and ECL2 (Fig. 4a; details

251   are given in Supplementary information, Table S10). Compared to the peptidic agonists

252   discussed above, the conformational architecture of THIQ reassembled the His-Phe-

253   Arg-Trp pharmacophore of α-MSH. Moiety R1 took the place of Trp9 of the peptide

254   agonist (Fig. 4a, b). The cyclohexane group from R1 fitted to a hydrophobic pocket

255   consisting of $S188^{ECL2}$, $I185^{4.61}$, $I194^{5.40}$ and $L197^{5.43}$, whereas triazole from R1 formed

256   a tight hydrogen bond with $H264^{6.54}$ and Van der Waals interactions with $I194^{5.40}$,

257     $L265^{6.55}$ and $Y268^{6.58}$ (Fig. 4a, b). Consistently, substitution of $H264^{6.54}$ with alanine

258   abolished the THIQ-induced activation (Fig. 4c). Moiety R2 D-Phe (pCl) structurally

259   mimicked Phe7, pointing to R1 and inserting into a hydrophobic pocket comprising

260     $I129^{3.32}$, $C130^{3.33}$, $L133^{3.36}$, $F261^{6.51}$ and $F284^{7.35}$. Moiety R3 plugged into another

261   hydrophobic pocket in which His6 of peptidic agonists was accommodated by $F45^{N\text{-}}$

262   ^{term}, $I104^{2.64}$ and $L288^{7.39}$, and was furthermore in close proximity to $N285^{7.36}$ (Fig. 4a,

263   b). However, similar to afamelanotide and bremelanotide, R3 of THIQ did not form a

264   hydrogen bond with $N285^{7.36}$ (Fig. 4a, b), whose replacement with alanine therefore

265   did not affect the potency of THIQ (Fig. 4c; details are given in Supplementary

266   information, Table S11). In contrast with peptidic agonists, only one carbonyl oxygen

267   atom from THIQ, equivalent to that of Phe7 in endogenous peptide ligand, participated

268   in coordinating Ca^{2+} (Fig. 4a). Hence, it appears that the function of THIQ is less

269   dependent on the metal ion when compared to α-MSH, as shown by our cAMP assay

270   (Supplementary information, Fig. S5e, h; details are given in Supplementary

271   information, Table S12).

272   In addition, we also docked several other small-molecule agonists into MC4R to

273   characterize the general activation mechanism of MC4R induced by non-peptidic

274   compounds. The docking poses of CHEMBL406764, MB-243 and RY764 highly

275   resembled the binding pose of THIQ, in which the p-chlorophenyl group of these

276   compounds inserted into the same hydrophobic pocket of MC4R as THIQ

277   (Supplementary information, Fig. S6a-f). Particularly, all of these three compounds

278   shared the same Ca^{2+} coordination, consistent with our observation of the THIQ-MC4R

279   complex structure (Supplementary information, Fig. S5d). However, the docking poses

280   indicate that moiety R3 was more flexible compared to moieties R1 and R2, in which

281   CHEMBL406764 is nearly perpendicular to THIQ at moiety R3 (Supplementary

282   information, Fig. S6d-f). This is probably due to the fact that the pockets for moieties

283   R1 and R2 are much deeper and more specific so that THIQ has relatively more

284   extensive interactions in these two pockets, whereas the pocket for moiety R3 is

285   shallow and flat in which THIQ has nonspecific hydrophobic interactions with only three

286   residues, *i.e.*, F45^{N-term}, I104$^{2.64}$ and L288$^{7.39}$. Nevertheless, the structure of the THIQ-

287   MC4R complex provides a template to understand the structure-activity relationship

288   (SAR) of small molecules targeting MC4R and facilitates the rational design of drugs

289   with the desired activity.

290

Receptor subtype selectivity

292   Structure-based sequence alignment suggests most of the residues in the peptidic

293   ligand binding pocket are highly conserved among MCR family members. Moreover,

294   critical residues for ligand binding and functional activity are completely conserved,

295   such as E100$^{2.60}$, D122$^{3.25}$, D126$^{3.29}$, I185$^{4.61}$, F261$^{6.51}$, H264$^{6.54}$, L265$^{6.55}$, F284$^{7.35}$ and

296   L288$^{7.39}$ (Supplementary information, Fig. S7a), providing the molecular basis of

297   polypharmacology of peptidic agonists of MCRs. Among the residues involved in

298   selective THIQ recognition, three are non-conserved between MC1R and MC4R, *i.e.*,

299   I129$^{3.32}$ at TM3, S188 at ECL2 and Y268$^{6.58}$ at the extracellular end of TM6, whereas

300   T124$^{3.32}$, Y183^{ECL2} and I264$^{6.58}$ are present in the corresponding positions in MC1R

301   (Fig. 5a, d). I129$^{3.32}$ is involved in forming the hydrophobic pocket accommodating R2

302   group of THIQ, whereas Y268$^{6.58}$ may interact with the imidazole ring. The substitution

303   of S188^{ECL2} with bulkier residues may hamper the entry of THIQ. We thus anticipate

304   that these three residues collaboratively determine the receptor subtype selectivity of

305   THIQ. To test our hypothesis, we replaced them with the corresponding residues of

306   MC1R and evaluated the potency of THIQ and α-MSH using the cAMP assay. It was

307   found that substitution of non-conserved residues dramatically reduced the potency of

308 THIQ by 5-20-fold (Fig. 5b; details are shown in Supplementary information, Table S11).

309 The combination of S188^{ECL2}Y and Y268$^{6.58}$I mutations also significantly decreased the

310 potency of THIQ over 50-fold (Fig. 5b, details are shown in Supplementary information,

311 Table S11). Furthermore, when we replaced all three residues of MC4R with those of

312 MC1R, THIQ-induced cAMP production was nearly eliminated (Fig. 5b, details are

313 shown in Supplementary information, Table S11). In contrast, different mutation

314 combinations resulted in negligible to mild effects on the functional activity of α-MSH

315 (Fig. 5c; details are shown in Supplementary information, Table S4). This data shows

316 that I129$^{3.32}$, S188^{ECL2} and Y268$^{6.58}$ in MC4R are critical residues responsible for

317 selective recognition of THIQ by this receptor subtype.

318   In summary, the sequence alignment of MC4R and MC1R displays 19 residue

319 variants in the orthosteric binding pocket, including 15 residues exhibiting different

320 chemical properties and 4 residues with similar properties (Supplementary information,

321 Fig. S7b-d). Structurally, most variants are located at the periphery of the 7TM

322 extracellular side. However, we noticed that two variants, I129$^{3.32}$ and L133$^{3.36}$ (T124$^{3.32}$

323 and M128$^{3.36}$ in MC1R) reside at the bottom of the ligand binding pocket, where they

324 are in close proximity to the His-Phe-Arg-Trp pharmacophore. It is also noteworthy that

325 four variants in the N terminus (Q43, L44, F45 and V46), one in TM6 (Y268$^{6.58}$) and

326 three in ECL3 (V278, M281 and S282) are located close to each other in a restricted

327 region, making it a hotspot for ligand specificity (Supplementary information, Fig. S7b-

328 d).

329

Role of Ca2+ in ligand activity

331 Ca^{2+} is considered a necessary cofactor for ligand binding and activity.[43] To evaluate

332 the effect of Ca^{2+} on the recognition of different ligands by MC4R, we gradually added

333 EGTA to remove Ca^{2+} from the medium and evaluated the binding and potency of the

334 MC4R ligands in question (Supplementary information, Fig. S5e-h; details are given in

335 Supplementary information, Table S12). As expected, α-MSH exhibited a complete

336 dependency on Ca^{2+}, evinced by the fact that the supplement of 5 mM EGTA abolished

337 the activity of α-MSH (Supplementary information, Fig. S5e; details are given in

338   Supplementary information, Table S12). The cyclic peptide bremelanotide behaved

339   similarly in lacking cAMP production following the addition of EGTA at different

340   concentrations (Supplementary information, Fig. S5g; details are given in

341   Supplementary information, Table S9, 12). Surprisingly, afamelanotide and THIQ

342   preserved 80% and 40% maximal efficacy of ligand-induced cAMP signaling in the

343   presence of 5 mM EGTA, respectively (Supplementary information, Fig. S5f, h; details

344   are given in Supplementary information, Table S12). Of the three MC4R residues

345   involved in Ca^{2+} action ($E100^{2.60}$, $D122^{3.25}$ and $D126^{3.29}$), mutating any of them would

346   abolish the effect of α-MSH (Fig. 2a, c; details are given in Supplementary information,

347   Table S4). Substitution of $D126^{3.29}A$ abolished bremelanotide-stimulated cAMP

348   accumulation, but only decreased the activities of afamelanotide and THIQ about 40-

349   and 100-fold, respectively. Similarly, mutant $E100^{2.60}A$ reduced the potencies of

350   bremelanotide and THIQ about 170- and 20-fold, respectively, but had almost no effect

351   on afamelanotide-induced MC4R activation (Fig. 3d, e, 4c; details are given in

352   Supplementary information, Table S4, 8, 9, 12). It thus appears that the activity of these

353   ligands is dependent on Ca^{2+} to varying degrees, raising the question of whether or

354   not there is a Ca^{2+}-independent mechanism for MC4R modulation.

355

Activation mechanism of MC4R

357   In comparison with the recently reported structure of SHU9119-bound tsMC4R in the

358   inactive state (PDB: 6W25),[43] binding to an agonist did not induce notable

359   conformational changes at the extracellular side of MC4R (Fig. 6a-b). This finding is in

360   contrast to previous structural studies of class A GPCRs, in which agonist-induced

361   receptor activation requires the extensive contraction of the 7TM extracellular half

362   correlated with the opening of the receptor cytoplasmic side to create a cavity for G

363   protein coupling.[36] Therefore, the most profound conformational changes of MC4R

364   upon activation were observed in the cytoplasmic regions that were stabilized by G_s

365   engagement, including (i) TM6 intracellular end exhibited 9.8 Å outward movement

366   measured at Cα of $K242^{6.32}$; and (ii) 6.9 Å inward movement of TM5 measured at Cα

367   of $L217^{5.63}$ (Fig. 6a-e).

368   Displacement of D-Phe7 in bremelanotide by the bulkier hydrophobic group D-

369   naphthylphenylalanine (D-Nal) led to SHU9119 and this subtle change sufficiently

370   turned a potent MC4R agonist into an antagonist (Fig. 6d; Supplementary information,

371   Fig. S8a-f, Table S2).[23] Therefore, peptidic agonists including α-MSH and SHU9119

372   shared a similar binding mode with MC4R, providing an explanation for the observation

373   that negligible conformational changes at the extracellular side were observed upon

374   receptor activation. The most profound conformational changes among the ligand

375   interacting residues occurred at the position L133$^{3.36}$. L/D-Phe7 of the agonists is 2 Å

376   shorter than that of D-Nal7 in SHU9119 when inserted into the binding pocket, leading

377   to the upward movement of L133$^{3.36}$ followed by a 2.9 Å downward movement of toggle

378   switch W258$^{6.48}$ (Fig. 6d). Substitution of L133$^{3.36}$ with alanine resulted in a dramatic

379   decrease in basal activity of the receptor, highlighting the important role of L133$^{3.36}$A

380   as a sensor defining the nature of the ligand during receptor activation (Supplementary

381   information, Fig. S8g). To support this hypothesis, we examined the SAR of peptide

382   agonist derivatives with different modifications of D-Phe7.[23,45] Apparently, D-Phe7, D-

383   Phe7(*p*F, *p* indicating a substituent in the *para* position of phenyl group) and D-

384   Phe7(*p*Cl) in either peptidic or small-molecule agonists displayed a similar optimal

385   potency, whereas D-Phe7(*p*I) caused a considerable steric hindrance with L133$^{3.36}$ due

386   to its larger Van der Walls radius reducing the efficacy by 33 folds compared to D-Phe7,

387   D-Phe7(*p*F) and D-Phe7(*p*Cl) (Supplementary information, Fig. S8a-f). The bulky D-

388   Nal group created a severe steric clash with L133$^{3.36}$ in the active state, and thus

389   pushed the side chain of L133$^{3.36}$ to rotate downwards facing the cytoplasmic side

390   accompanied by an upward movement of W258$^{6.48}$, resulting in the "toggle twin switch"

391   (L133$^{3.36}$ and W258$^{6.48}$) in a canonical inactive conformation of class A GPCR (Fig. 6d).

392   On the other hand, swapping NDP-MSH residue D-Phe7 to D-Ala7 would result in the

393   loss of the hydrophobic interaction between the phenyl group and L133$^{3.36}$, thereby

394   insufficiently stabilizing L133$^{3.36}$ in the upward active conformation (Supplementary

395   information, Fig. S8a). Consequently, this substitution profoundly decreased ligand

396   potency by over 1000-fold.[39] Markedly, substituting the residue L133$^{3.36}$ to methionine

397   led to the complete conversion of SHU9119 activity from antagonist to agonist of

13

398   MC4R,[46] possibly owing to the fact that the methionine side chain has more freedom

399   to adopt different rotamers accommodating the bulky D-Nal7 residue of SHU9119,

400   further emphasizing the critical role of L133$^{3.36}$ in receptor activation. In conclusion, our

401   agonist-bound structures uncovered the MC4R agonism which is precisely modulated

402   by the spatial size of the residue located at the peptide 7-position or the equivalent

403   moiety of small-molecule compounds (Supplementary information, Fig. S8a-f).

404   The concerted conformational changes of the "toggle twin switch" L133$^{3.36}$ and

405   W258$^{6.48}$ relayed by M204$^{5.50}$I137$^{3.40}$F254$^{6.44}$ motif in the MC receptor family, are

406   equivalent to the conserved PIF motif in other class A GPCRs.[47] The above-mentioned

407   agonist-induced structural changes drove W258$^{6.48}$ to slide downward contacting

408   L133$^{3.36}$ and I137$^{3.40}$ propagated through F254$^{6.44}$ (Fig. 6d). The outward movement of

409   F254$^{6.44}$ opened up TM6 by breaking the hydrophobic interaction between L140$^{3.43}$ and

410   I143$^{3.46}$ in TM3 and I251$^{6.41}$ and L247$^{6.37}$ in TM6 and the hydrogen bond between

411   R147$^{3.50}$ and N240$^{6.30}$, creating a cytoplasmic cavity for α5 of Gα$_s$ (Fig. 6e). This

412   suggests that F254$^{6.44}$ also plays a key role in MC4R activation

413   Interestingly, TM5 exhibited a profound yet unique conformational change induced

414   by agonist and G protein binding (Fig. 6f). The structure superposition showed that

415   TM5 extended three helical turns and moved inward 6.9 Å measured at the Cα atom

416   of L217$^{5.63}$ upon activation (Fig. 6a, f). The lipid separating TM3 and TM5 in the inactive

417   MC4R was squeezed off from the activated receptor as we did not observe any density

418   at the equivalent position (Supplementary information, Fig. S8h-j). The interactions

419   between TM3 and TM5 were fastened by the newly established hydrogen bond

420   between Y148$^{3.51}$ and H214$^{5.60}$. The hydrogen bond between H214$^{5.60}$ and H222$^{5.68}$ in

421   the inactive state was broken by the extension of TM5 (Supplementary information,

422   Fig. S8h-j). In contrast, previous structural studies of class A GPCRs, such as β$_2$AR,

423   A$_{2A}$R, β$_1$AR and EP$_4$, demonstrated that receptor activation always induces an outward

424   movement of TM5 to various extents, 1.4-2.8 Å measured at Cα of the residue 5.66

425   (β$_2$AR: K277; A$_{2A}$R: R205; β$_1$AR: K235; and EP$_4$: L211) (Fig. 6f).[48-55] The distinctive

426   structural rearrangement of TM5 related to MC4R activation probably attributes to the

427   presence of a lipid between TM3 and TM5 in addition to the absence of the conserved

14

428   Pro at 5.50.[47]

429

MC4R-G$_s$ coupling interface

431   The MC4R-G$_s$ interface involves TM3, TMs 5-7 and all three ICLs (Fig. 7; details are

432   shown in Supplementary information, Table S13). The overall cryo-EM structures of

433   the MC4R-G$_s$ complexes revealed a similar interaction mode of the MC4R-G protein

434   compared to other class A GPCRs, such as β$_2$AR-G$_s$ (PDB: 3SN6), A$_{2A}$R-G$_s$ (PDB:

435   6GDG) and GPBAR-G$_s$ (PDB: 7CFM) [34,49,56] (Supplementary information, Fig. S9a-b).

436   However, the high-resolution density map showed several distinctive features of the

437   MC4R-G$_s$ complex. In other reported class A GPCR-G$_s$ complexes, the side chain of

438   Y391$^{G.H5.23}$ in Gα helix 5 was recognized primarily by packing against TM3 in parallel

439   with R$^{3.50}$ from the conserved DRY motif with the distance of 4 Å between two side

440   chain planes and did not established strong interactions with the receptors. In the

441   MC4R-G$_s$ structures, however, the hydroxyl group of Y391$^{G.H5.23}$ formed a tight

442   hydrogen bond with the side chain of T150$^{3.53}$, a phenomenon strictly conserved in the

443   MCR family. Residues at the equivalent position in other class A GPCRs with available

444   structures in complex with G$_s$ are alanine (Fig. 7f; Supplementary information, Fig. S9c-

445   d). Strikingly, displacement of T150$^{3.53}$ with alanine completely abolished the agonist-

446   induced cAMP production but did not affect the β-arrestin recruitment to MC4R,

447   suggesting that T150$^{3.53}$ plays an important role in G$_s$ coupling to MC4R but may not

448   be involved in β-arrestin signaling (Fig. 7g-h; details are shown in Supplementary

449   information, Table S14).

450   Of note, we observed multiple well-defined water molecules participating in the

451   Gα$_s$ engagement to MC4R. To our knowledge, this is the first description of water

452   molecule involvement in G protein coupling among class A GPCRs. It appears that

453   water molecules connect three pairs of residues in the interface, E392$^{G.H5.24}$ to R147$^{3.50}$,

454   H387$^{G.H5.19}$ to Y157^{ICL2} and H41$^{G.S1.02}$ to H158^{ICL2} (Fig. 7a). E392$^{G.H5.24}$ also forms a salt

455   bridge with R305$^{7.56}$ in the TM7-H8 hinge, similar to other class A GPCRs (Fig. 7b). We

456   individually mutated these residues in the G$_s$ pocket, and assessed their effects on

457   expression level and ability to stimulate a cAMP response. The results showed that

15

458   R147$^{3.50}$A, Y157^{ICL2}A and R305$^{7.56}$A reduced ligand potency 5-20-fold, pointing to

459   crucial roles of these residues in G$_s$ signaling (Fig. 7e; details are given in

460   Supplementary information, Table S4). Though all the ICLs are involved in the Gs

461   coupling, our mutagenesis study suggested that ICL2 played a more critical role

462   compared to ICL1 and ICL3, supported by the fact substitutions of L155^{ICL2}A and

463   Y157^{ICL2}A decreased α-MSH-induced receptor activation 10-20-fold whereas the

464   mutations of residues in ICL1or ICL3 barely affected the ligand activity (Fig. 7e; details

465   are given in Supplementary information, Table S4).

466

467   ## Naturally occurring mutations in MC4R

468   To date, more than 200 genetic variants have been discovered in the MC4R gene [57].

469   Human genetic studies show that loss-of-function (LoF) mutations in the MC4R gene

470   are the most common monogenic cause of obesity. However, recent genetic

471   association studies of approximately 0.5 million samples from the UK Biobank indicate

472   that a subset of MC4R variants can cause gain-of-function (GoF) phenotypes, such as

473   G231^{ICL3}S, F201$^{5.47}$L, I251$^{6.41}$L and L304$^{7.55}$F. Especially, GoF genetic variants in

474   MC4R exhibited a signal bias for β-arrestin recruitment over canonical Gα$_s$-mediated

475   cAMP production and may have implications in obesity prevention.[8] Therefore,

476   structural determination of the activated MC4R-G$_s$ complex offers a great opportunity

477   to examine the molecular mechanism of functionally defective or enhanced variants.

478   Mutation of G231^{ICL3} to serine could potentially form a hydrogen bond with the

479   spatially neighboring residue T350$^{G.h4S6.03}$ of Gα resulting in enhanced Gα$_s$-MC4R

480   interaction, thus perfectly explaining the increased Gα$_s$-mediated cAMP production by

481   G231^{ICL3}S mutation.[8] However, another natural mutation at this position G231^{ICL3}V that

482   lacks the hydrogen bond with T350$^{G.h4S6.03}$ did not affect G$_s$ signaling.[8] Interestingly,

483   G231^{ICL3}V/S variants in MC4R did not affect β-arrestin recruitment, indicating that

484   G231^{ICL3} may not participate in β-arrestin signaling. In close proximity to the critical

485   microswitch "toggle switch" and P/MIF motif, F201$^{5.47}$L and I251$^{6.41}$L variants may

486   stabilize the receptor in the active state, leading to higher efficacies. At the cytoplasmic

487   side, the hinge betweenTM7-H8 connects these two helices to TM1-ICL1-TM2, and

488   plays important roles in G protein coupling and β-arrestin recruitment. Mutation of

489   L304[7.55] to phenylalanine with a larger side chain, which potentially increases the

490   conformational rigidity of TM7-TM8-TM1-TM2 region by forming more hydrophobic

491   interactions with the surrounding residues including L309[8.50], R310[8.51] and F313[8.54],

492   increased cAMP production and β-arrestin recruitment.[8]

493   A classification system was proposed to sort obesity-associated MC4R

494   mutations.[58] Many mutations that were functionally characterized as primarily defective

495   in trafficking (class II), or with no apparent defect (class V) are actually involved in G-

496   protein coupling according to our structural information and functional experiments.

497   For example, T150[3.53], Y157[ICL2] and R305[7.56] (class II) as well as Q156[ICL2] and T162[4.38]

498   (class V) were found at the G-protein interface, and mutagenesis data show that these

499   residues are crucial for G_s coupling (Fig.7d). Therefore, these mutants should also be

500   included in class IV (defective in coupling to $G\alpha_s$). Additionally, functional results

501   indicate that I317[8.58] mutants (class IV) did not affect G protein signaling,[8] consistent

502   with our structural observation that I317[8.58] does not obviously interact with the G

503   protein. Thus, it should not be classified as class IV (defective coupling to $G\alpha_s$) mutants.

504   [52] Naturally occurring mutants are closely related to obesity, and the detailed functional

505   studies and accurate classifications are of great significance for practicing precision

506   medicine.

507

508   **DISCUSSION**

509   Delineating the structural basis for ligand recognition and G protein coupling at MC4R

510   will aid in the rational design of drugs with high selectivity and reduced side effects. It

511   will also facilitate our knowledge about naturally occurring MC4R mutants in the

512   context of disease treatments. Here, we reported high-resolution cryo-EM structures

513   of MC4R signaling complexes bound to four agonists with distinct chemical and

514   functional features, including three non-selective peptidic agonists, *i.e.* endogenous

515   peptide α-MSH, linear peptide afamelanotide and cyclic peptide bremelanotide, and

516   one MC4R-selective small-molecule compound THIQ. Both peptide agonists and THIQ

517   bind to the center of the wild-open binding pocket in the extracellular half of the MC4R

518   7TM bundle. The "toggle twin switch" L133$^{3.36}$ and W258$^{6.48}$, packed against the ligand

519   phenyl ring, sense ligand binding and determine the functional nature of ligands. These

520   structures together with mutagenesis studies revealed a common recognition mode of

521   peptide agonists of the MCR family as well as critical determinants including TM3

522   residue I129$^{3.32}$, ECL2 residue S188 and TM6 residue Y268$^{6.58}$ responsible for THIQ

523   selectivity. Furthermore, the structures also disclosed the MC4R-G$_s$ engaging interface,

524   providing the molecular basis for G protein-coupling specificity. In addition, all three

525   ICLs of MC4R are involved in direct interactions with the G$_s$ trimer, which may

526   contribute to the high basal activity of MC4R. Together, our results offer unprecedented

527   structural insights into the pharmacological and signaling features of MC4R and

528   deepen our understanding of how naturally occurring mutants of MC4R modify the

529   receptor activity. Our findings will serve as a structural template for rational drug design

530   targeting the leptin-melanocortin pathway and facilitate the discovery of novel

531   therapeutics against obesity. Moreover, these observed characteristics in the MC4R

532   structures, including structural arrangement of the 7TM bundle, ligand-binding mode,

533   agonist-induced receptor activation, and the manner of G protein coupling, could also

534   be shared by other MCR family members.

535

536   **METHODS**

537   **Cloning and insect cell expression**

538   The wild-type (WT) human melanocortin-4 receptor (UniProt ID: P32245) gene was

539   cloned into a modified pFastBac1 vector with hemagglutinin (HA) signal sequence at

540   the N terminus and a PreScission protease site followed by a Flag tag. LgBiT and a

541   double MBP tag were fused to the C terminus with a 3C protease cleavage site

542   between them. A dominant-negative human Gα$_s$ (DNGα$_s$) was generated by site-

543   directed mutagenesis as previously described to stabilize the interaction with the βγ

544   subunits [31,59] and cloned into pFastaBac1 vector. Gβ1 was fused with an N-terminal

545   HiBit and 10 × His tag, together with Gγ2 were cloned into pFastBac dual vector.

546   The WT MC4R, DNGα$_s$ and Gβ1γ2 constructs were co-expressed in *Spodoptera*

547   *frugiperda* (*Sf*9) insect cells using the Bac-to-Bac Baculovirus Expression System

548   (Invitrogen). Cells were infected at a density of 2.4 × 10^6 cells per mL and then co-

549   infected with three separate viruses at a ratio of 1:1:1 for MC4R, DNGα$_s$ and Gβ1γ2.

550   Cells were collected after 48 h post-infection and stored at -80°C until use.

Expression and purification of Nb35

552   Nb35 with a C-terminal 6 × His-tag was expressed in the periplasm of *Escherichia. Coli*

553   WK6 cells. The cells containing the recombinant plasmid were cultured in TB media

554   supplemented with 0.1% glucose, 1 mM MgCl$_2$ and 50μg/mL kanamycin at 37°C until

555   OD600 reached 0.6. Then the cultures were induced by 1 mM IPTG and grown at 18°C

556   for 24 h. The cells were harvested by centrifugation at 4,000 *g* for 15 min, and

557   subsequently lysed. The protein was purified by nickel affinity chromatography as

558   previously described [34]. Eluted protein was concentrated using a 10 kDa molecular

559   weight cut-off concentrator (Millipore) and loaded onto a HiLoad 16/600 Superdex 75

560   column (GE Healthcare) with running buffer containing 20 mM HEPES pH 7.5 and 100

561   mM NaCl. The monomeric fractions were pooled. Purified Nb35 was finally flash frozen

562   in liquid nitrogen and stored at -80°C.

Purification of MC4R–G$_s$–Nb35 complexes

564   The cell pellets were thawed on ice and lysed in a buffer containing 20 mM HEPES,

565   pH 7.5, 100 mM NaCl and 2 mM MgCl$_2$ supplemented with EDTA-free protease

566   inhibitor cocktail (Bimake) by dounce homogenization. The complex formation was

567   initiated by addition of 10 μg/mL Nb35, 25 mU/mL apyrase (NEB) and adequate agonist

568   (100 μM α-MSH, 100 μM afamelanotide, 20 μM bremelanotide; China Peptides) or 10

569   μM THIQ (TopScience). The cell lysate was subsequently incubated for 1 h at room

570   temperature (RT) and then solubilized by 0.5% (w/v) lauryl maltose neopentyl glycol

571   (LMNG, Anatrace) and 0.1% (w/v) cholesterol hemisuccinate (CHS, Anatrace) for 2 h

572   at 4ºC. After centrifugation at 30,000 *g* for 30 min, the supernatant was isolated and

573   incubated with amylose resin (NEB) for 2 h at 4ºC. Then the resin was collected by

574   centrifugation at 600 *g* for 10 min and loaded into a gravity flow column (Beyotime),

575   and washed with five column volumes (CVs) of buffer containing 20 mM HEPES, pH

576   7.5, 100 mM NaCl, 2 mM MgCl$_2$, agonist (100 μM α-MSH, 100 μM afamelanotide, 20

577   μM bremelanotide or 10 μM THIQ), 0.01% (w/v) LMNG and 0.005% (w/v) CHS, eluted

578   with 15 CVs of buffer containing 20 mM HEPES, pH 7.5, 100 mM NaCl, 2 mM MgCl$_2$,

579   agonist (100 µM α-MSH, 10 µM afamelanotide, 20 µM bremelanotide or 10 µM THIQ),

580   0.01% (w/v) LMNG, 0.005% (w/v) CHS and 10 mM maltose. The elution was collected

581   and incubated with 3C protease for 1 h at RT. Then the elution was concentrated with

582   a 100 kDa cut-off concentrator (Millipore). Concentrated complex was loaded onto a

583   Superose 6 increase 10/300 GL column (GE Healthcare) with running buffer containing

584   20 mM HEPES, pH 7.5, 100 mM NaCl, 2 mM MgCl$_2$, agonist (100 µM α-MSH, 10 µM

585   afamelanotide, 20 µM bremelanotide or 10 µM THIQ), 0.00075% (w/v) LMNG, 0.0002%

586   (w/v) CHS and 0.00025% (w/v) GDN (Anatrace). The fractions for the monomeric

587   complex were collected and concentrated for electron microscopy experiments.

588   **Cryo-EM grid preparation and data collection**

589   For cryo-EM grid preparation, 3 µL of the purified α-MSH–, afamelanotide–,

590   bremelanotide– or THIQ–bound MC4R–G$_s$ complexes at ~ 4 mg/mL were applied onto

591   glow-discharged holey carbon grids (Quantifoil R1.2/1.3). Excess samples were

592   blotted away and the grids were vitrified by plunging into liquid ethane using a Vitrobot

593   Mark IV (ThermoFischer Scientific).

594   Cryo-EM imaging was performed on a Titan Krios at 300 kV in the Center of Cryo-

595   Electron Microscopy, Zhejiang University (Hangzhou, China). Micrographs were

596   recorded using a Gatan K2 Summit detector in counting mode with a pixel size of 1.014

597   Å using the SerialEM software. Movies were obtained at a dose rate of about 7.8 e/Å^2/s

598   with a defocus ranging from -0.5 to -2.0 µm. The exposure time was 8 s and 40 frames

599   were recorded per micrograph. A total of 3,321, 3,521, 3,018 and 5,731 movies were

600   collected for the α-MSH–, afamelanotide–, bremelanotide– and THIQ–bound MC4R–

601   G$_s$ complexes, respectively.

602   **Cryo-EM data processing**

603   Image stacks were aligned using MotionCor 2.1.[60] Contrast transfer function (CTF)

604   parameters were estimated by Gctf v1.18.[61] The following data processing was

605   performed using RELION-3.0-beta2.[62]

606   For the α-MSH–bound MC4R–G$_s$ complex, automated particle selection using

607   Gaussian blob detection produced 2,170,482 particles. The particles were subjected

608   to reference-free 2D classification to discard fuzzy particles, resulting in 680,962

609   particles for further processing. The map of GPBAR–G_s complex (EMD-30344)[56] low-

610   pass filtered to 60 Å was used as the reference map for 3D classification, generating

611   one well-defined subset with 167,001 particles. Further 3D classifications focusing the

612   alignment on the complex, produced three good subsets accounting for 147,565

613   particles, which were subsequently subjected to 3D refinement, CTF refinement and

614   Bayesian polishing. The final refinement generated a map with an indicated global

615   resolution of 3.0 Å at a Fourier shell correlation of 0.143.

616   For the afamelanotide–bound complex, 2,179,940 particles generated from the

617   automated particle picking were subjected to 2D classification, producing 1,137,998

618   particles for 3D classification. The map of the α-MSH–bound complex was used as the

619   reference for initial 3D classification, producing one good subset with 470,528 particles.

620   Further 2 rounds of 3D classifications focusing the alignment on the MC4R–G_s

621   complex or the MC4R receptor produced two high-quality subsets accounting for

622   217,491 particles, which were subsequently subjected to 3D refinement, CTF

623   refinement and Bayesian polishing. The final refinement generated a map with an

624   indicated global resolution of 3.0 Å.

625   For the bremelanotide–bound complex, automated particle picking yielded

626   2,259,228 particles, which were subjected to 2D classification. The well-defined

627   classes were selected, producing 902,777 particles for further processing. The density

628   map of α-MSH-bound complex low-pass filtered to 40 Å was used as the reference

629   map for 3D classification, producing two good subsets with 274,222 particles. The

630   particles were subsequently subjected to 3D refinement, CTF refinement and Bayesian

631   polishing. The final refinement generated a map with an indicated global resolution of

632   3.1 Å.

633   For the THIQ–bound complex, automated particle picking yielded 2,905,757

634   particles, which were subjected to 2D classification. The well-defined classes were

635   selected, producing 1,148,249 particles for further processing. The density map of α-

636   MSH-bound complex low-pass filtered to 40 Å was used as the reference map for 3D

637   classification, producing two good subsets with 269,832 particles. The particles were

21

638   subsequently subjected to 3D refinement, CTF refinement and Bayesian polishing.

639   The final refinement generated a map with an indicated global resolution of 2.9 Å.

640   Local resolution was determined using the Bsoft package with half maps as input

641   maps.[63]

Model building, refinement and validation

643   The initial homology model of the active state MC4R was generated from GPCRdb.[42]

644   The parathyroid hormone receptor-1–G_s–Nb35 complex (PDB 6NBF)[64] was to

645   generate the initial models of G_s and Nb35. Small molecular ligand coordinates and

646   geometry restraints were generated using phenix.elbow. Models were docked into the

647   EM density map using UCSF Chimera (https://www.cgl.ucsf.edu/chimera/). This

648   starting model was then subjected to iterative rounds of manual adjustment and

649   automated refinement in Coot[65] and Phenix[66], respectively.

650   The final refinement statistics were validated using the module 'comprehensive

651   validation (cryo-EM)' in PHENIX. To monitor the potential over-fitting in model building,

652   FSC_{work} and FSC_{test} were determined by refining the 'shaken' models against unfiltered

653   half-map-1 and calculating the FSC of the refined models against unfiltered half-map-

654   1 and half-map-2. Structural figures were prepared in Chimera

655   (https://www.cgl.ucsf.edu/chimera), ChimeraX[67,68] and PyMOL (https://pymol.org/2/).

656   The final refinement statistics are provided in Supplementary Table S2.

cAMP accumulation assay

658   Agonist (α-MSH, afamelanotide, bremelanotide or THIQ) stimulated cAMP

659   accumulation was measured by a GloSensor™ cAMP assay kit (Promega). Briefly,

660   HEK293T cells transfected with WT or mutant MC4R and the pGloSensor™-22F

661   plasmid were seeded onto 384-well culture plates at a density of 4×10^3 cells per well

662   and incubated for 24 h at 37°C in 5% CO_2. Then the culture medium was removed and

663   the equilibration medium containing 4% (v/v) dilution of the GloSensor™ cAMP reagent

664   stock solution was added to each well. To analyze the effect of Ca^{2+} on cAMP signaling,

665   the equilibration medium was further supplemented with different concentrations of

666   EGTA (0 mM, 1 mM, 5 mM, 10 mM and 20 mM). To obtain the concentration-response

667   curves, serially diluted agonists were added to each well to stimulate the cells.

22

668   Luminance signal was measured using 0.5 s intervals after ligand addition (TECAN,

669   25°C). Concentration-responses were generated from the peak response. cAMP

670   accumulation was analyzed by a standard dose-response curve using GraphPad

671   Prism 8.0 (GraphPad Software). EC_{50} and pEC_{50} ± SEM were calculated using

672   nonlinear regression (curve fit). Data are means ± SEM from at least three independent

673   experiments performed in technical triplicates.

Receptor expression

675   Plasmids corresponding to WT and mutant MC4R, or pcDNA3.1 (+) empty vector, were

676   transfected as described above. After transfection, cells were re-seeded onto poly-D-

677   lysine treated 96-well plates at a density of 3×10^4 cells per well. Twenty-four hours

678   later, cells were washed with PBS and fixed with 10% formaldehyde for 10 min followed

679   by three times washing with PBS. Following fixation, cells were blocked with blocking

680   buffer (1% BSA in PBS) for 1 h at RT. Afterward, plates were incubated with a 1:20,000

681   dilution of anti-FLAG M2 HRP-conjugated monoclonal antibody (Sigma-Aldrich) in

682   blocking buffer for another 1 h at RT. After careful washing, 80 µL/well diluent

683   SuperSignal Elisa Femto Maximum Sensitivity Substrate (ThermoFisher Scientific)

684   was added. Luminance signal was measured using 1 s intervals. Data are means ±

685   SEM from at least three independent experiments performed in technical triplicates.

Radioligand binding assay

687   Radioligand binding assays for WT and mutant MC4R were performed in transiently

688   transfected HEK293 cells. Briefly, cells were seeded at a density of 3×10^4 cells/well

689   onto 96-well culture plates and incubated overnight at 37°C in 5% CO_2, and

690   competitive radioligand binding was assessed 24 h thereafter. For homogeneous

691   binding, the cells were incubated in binding buffer (DMEM supplemented with 25 mM

692   HEPES and 0.1% BSA) with a constant concentration of ^{125}I-α-MSH, (40 pM,

693   PerkinElmer) and different concentrations of unlabeled peptides [α-MSH (47.68 pM to

694   100 µM); afamelanotide (1 pM to 10 µM); bremelanotide (1 pM to 10 µM)] at RT for 3

695   h. Then plates were washed three times with ice-cold PBS and lysed by 50 µl lysis

696   buffer (PBS supplemented with 20 mM Tris-HCl, pH 7.4 and 1% Triton X-100). The

697   plates were subsequently counted for radioactivity (counts per minute, CPM) with a

698   MicroBeta^2 plate counter (PerkinElmer) using a scintillation cocktail (OptiPhase

699   SuperMix, PerkinElmer).

700   **β-arrestin 2 recruitment assay**

701   β-arrestin 2 recruitment was measured in HEK293T cells using a bioluminescence

702   resonance energy transfer (BRET) assay. The cells were seeded onto poly-D-lysine

703   coated 96-well plates at a density of 3.5×10^4 cells/well and grown overnight before

704   transfection. Prior to BRET experiments, cells were transiently transfected with MC4R-

705   Rluc8 and β-arrestin 2-Venus at a ratio of 1:6 for 24 h and rinsed twice with HBSS. The

706   cells were then incubated with BRET buffer (HBSS supplemented with 10 mM HEPES

707   and 0.1% BSA, pH 7.4) for 30 min at 37°C. After incubation with 5 μM coelenterazine-

708   H (Yeasen Biotech) for 5 min, the baseline BRET signals were read immediately at 470

709   nm and 535 nm for 10 cycles using an EnVision® multimode plate reader (PerkinElmer).

710   Following the agonist addition, BRET was measured every 1 min for another 50 cycles.

711   Data are presented as BRET ratio, calculated as the ratio of Venus to Rluc8 signals

712   and subtracted the vehicle value. Concentration-response profiles were obtained from

713   area-under-the-curve values of elicited responses.

714   **Molecular docking**

715   The active MC4R model was generated from the high-resolution structure of α-MSH–

716   MC4R–G$_s$ complex for molecular docking. Three representative analogs of small

717   molecule THIQ were docked into the active MC4R model using the "RosettaLigand"

718   application in Rosetta 2019.35 version.[69,70] Conformers of the three ligands were

719   generated by the Open Babel program (The Open Babel Package, version 2.3.1

720   http://openbabel.org).[71] Then the ligand docking poses were searched from the pocket

721   of MC4R with a box size of 15 Å controlled by the RosettaScrpits. A total of 1,000

722   models were generated for the docking trial. The top 10 docking poses of each ligand

723   were selected based on the lowest binding energy (interface_delta_X) and viewed by

724   PyMol. The most similar binding pose as THIQ of each ligand was selected and

725   displayed in Supplementary Figure S6.

726   **Statistical analysis**

727   All functional study data were analyzed using Prism 8 (GraphPad) and presented as

728   means ± S.E.M. from at least three independent experiments. Concentration-response

729   curves were evaluated with a standard dose-response curve. Statistical differences

730   were determined by two-sided, one-way ANOVA with Dunnett's multiple comparisons

731   test.

732   **Data availability**

733   The atomic coordinates and the electron microscopy maps of the α-MSH–,

734   afamelanotide–, bremelanotide– and THIQ–bound MC4R–G_s complexes have been

735   deposited in the Protein Data Bank (PDB) under accession number xxx, xxx, xxx, xxx

736   and Electron Microscopy Data Bank (EMDB) under accession codes xxx, xxx, xxx, xxx

737   respectively. All relevant data are included in the manuscript or Supplementary

738   Information.

739

740   **References**

741   1   Krashes, M. J., Lowell, B. B. & Garfield, A. S. Melanocortin-4 receptor-regulated energy
742   homeostasis. *Nature neuroscience* **19**, 206-219, doi:10.1038/nn.4202 (2016).

743   2   Tao, Y. X. The melanocortin-4 receptor: physiology, pharmacology, and pathophysiology.
744   *Endocrine reviews* **31**, 506-543, doi:10.1210/er.2009-0037 (2010).

745   3   Kuhnen, P., Krude, H. & Biebermann, H. Melanocortin-4 Receptor Signalling: Importance for
746   Weight Regulation and Obesity Treatment. *Trends in molecular medicine* **25**, 136-148,
747   doi:10.1016/j.molmed.2018.12.002 (2019).

748   4   Siljee, J. E. *et al.* Subcellular localization of MC4R with ADCY3 at neuronal primary cilia underlies
749   a common pathway for genetic predisposition to obesity. *Nature genetics* **50**, 180-185,
750   doi:10.1038/s41588-017-0020-9 (2018).

751   5   Hinney, A., Volckmar, A. L. & Knoll, N. Melanocortin-4 receptor in energy homeostasis and
752   obesity pathogenesis. *Progress in molecular biology and translational science* **114**, 147-191,
753   doi:10.1016/b978-0-12-386933-3.00005-4 (2013).

754   6   Yang, Y. Structure, function and regulation of the melanocortin receptors. *European journal of*
755   *pharmacology* **660**, 125-130, doi:10.1016/j.ejphar.2010.12.020 (2011).

756   7   Vaisse, C., Clement, K., Guy-Grand, B. & Froguel, P. A frameshift mutation in human MC4R is
757   associated with a dominant form of obesity. *Nature genetics* **20**, 113-114, doi:10.1038/2407
758   (1998).

759   8   Lotta, L. A. *et al.* Human Gain-of-Function MC4R Variants Show Signaling Bias and Protect
760   against Obesity. *Cell* **177**, 597-607.e599, doi:10.1016/j.cell.2019.03.044 (2019).

761   9   Branson, R. *et al.* Binge eating as a major phenotype of melanocortin 4 receptor gene
762   mutations. *The New England journal of medicine* **348**, 1096-1103, doi:10.1056/NEJMoa021971
763   (2003).

764   10   Collet, T. H. *et al.* Evaluation of a melanocortin-4 receptor (MC4R) agonist (Setmelanotide) in
765   MC4R deficiency. *Molecular metabolism* **6**, 1321-1329, doi:10.1016/j.molmet.2017.06.015

766   (2017).

767 11 Goncalves, J. P. L., Palmer, D. & Meldal, M. MC4R Agonists: Structural Overview on Antiobesity
768   Therapeutics. *Trends in pharmacological sciences* **39**, 402-423, doi:10.1016/j.tips.2018.01.004
769   (2018).

770 12 Cooray, S. N. & Clark, A. J. Melanocortin receptors and their accessory proteins. *Molecular and*
771   *cellular endocrinology* **331**, 215-221, doi:10.1016/j.mce.2010.07.015 (2011).

772 13 Gantz, I. *et al.* Molecular cloning of a novel melanocortin receptor. *The Journal of biological*
773   *chemistry* **268**, 8246-8250 (1993).

774 14 Ghamari-Langroudi, M. *et al.* Regulation of energy rheostasis by the melanocortin-3 receptor.
775   *Science advances* **4**, eaat0866, doi:10.1126/sciadv.aat0866 (2018).

776 15 van der Kraan, M. *et al.* Expression of melanocortin-5 receptor in secretory epithelia supports
777   a functional role in exocrine and endocrine glands. *Endocrinology* **139**, 2348-2355,
778   doi:10.1210/endo.139.5.6008 (1998).

779 16 Catania, A., Gatti, S., Colombo, G. & Lipton, J. M. Targeting melanocortin receptors as a novel
780   strategy to control inflammation. *Pharmacological reviews* **56**, 1-29, doi:10.1124/pr.56.1.1
781   (2004).

782 17 Cone, R. D. Studies on the physiological functions of the melanocortin system. *Endocrine*
783   *reviews* **27**, 736-749, doi:10.1210/er.2006-0034 (2006).

784 18 Metherell, L. A. *et al.* Mutations in MRAP, encoding a new interacting partner of the ACTH
785   receptor, cause familial glucocorticoid deficiency type 2. *Nature genetics* **37**, 166-170,
786   doi:10.1038/ng1501 (2005).

787 19 Dores, R. M. *et al.* Molecular evolution of GPCRs: Melanocortin/melanocortin receptors.
788   *Journal of molecular endocrinology* **52**, T29-42, doi:10.1530/jme-14-0050 (2014).

789 20 Nargund, R. P., Strack, A. M. & Fong, T. M. Melanocortin-4 receptor (MC4R) agonists for the
790   treatment of obesity. *Journal of medicinal chemistry* **49**, 4035-4043, doi:10.1021/jm058241a
791   (2006).

792 21 Han, G. *et al.* De novo design, synthesis, and pharmacology of alpha-melanocyte stimulating
793   hormone analogues derived from somatostatin by a hybrid approach. *Journal of medicinal*
794   *chemistry* **47**, 1514-1526, doi:10.1021/jm030452x (2004).

795 22 Al-Obeidi, F., Castrucci, A. M., Hadley, M. E. & Hruby, V. J. Potent and prolonged acting cyclic
796   lactam analogues of alpha-melanotropin: design based on molecular dynamics. *Journal of*
797   *medicinal chemistry* **32**, 2555-2561, doi:10.1021/jm00132a010 (1989).

798 23 Hruby, V. J. *et al.* Cyclic lactam alpha-melanotropin analogues of Ac-Nle4-cyclo[Asp5, D-
799   Phe7,Lys10] alpha-melanocyte-stimulating hormone-(4-10)-NH2 with bulky aromatic amino
800   acids at position 7 show high antagonist potency and selectivity at specific melanocortin
801   receptors. *Journal of medicinal chemistry* **38**, 3454-3461, doi:10.1021/jm00018a005 (1995).

802 24 Holder, J. R. & Haskell-Luevano, C. Melanocortin ligands: 30 years of structure-activity
803   relationship (SAR) studies. *Medicinal research reviews* **24**, 325-356, doi:10.1002/med.10064
804   (2004).

805 25 Sebhat, I. K. *et al.* Design and pharmacology of N-[(3R)-1,2,3,4-tetrahydroisoquinolinium- 3-
806   ylcarbonyl]-(1R)-1-(4-chlorobenzyl)- 2-[4-cyclohexyl-4-(1H-1,2,4-triazol- 1-ylmethyl)piperidin-
807   1-yl]-2-oxoethylamine (1), a potent, selective, melanocortin subtype-4 receptor agonist.
808   *Journal of medicinal chemistry* **45**, 4589-4593, doi:10.1021/jm025539h (2002).

809 26 Ye, Z. *et al.* Discovery and activity of (1R,4S,6R)-N-[(1R)-2-[4-cyclohexyl-4-[[(1,1-

810   dimethylethyl)amino]carbonyl]-1-piper   idinyl]-1-[(4-fluorophenyl)methyl]-2-oxoethyl]-2-

811   methyl-2-azabicyclo[2.2.2]octane -6-carboxamide (3, RY764), a potent and selective

812   melanocortin subtype-4 receptor agonist. *Bioorganic & medicinal chemistry letters* **15**, 3501-

813   3505, doi:10.1016/j.bmcl.2005.05.109 (2005).

814   27   Matharu, N. *et al.* CRISPR-mediated activation of a promoter or enhancer rescues obesity

815   caused by haploinsufficiency. *Science (New York, N.Y.)* **363**, doi:10.1126/science.aau0629

816   (2019).

817   28   Sawyer, T. K. *et al.* 4-Norleucine, 7-D-phenylalanine-alpha-melanocyte-stimulating hormone: a

818   highly potent alpha-melanotropin with ultralong biological activity. *Proceedings of the National*

819   *Academy of Sciences of the United States of America* **77**, 5754-5758,

820   doi:10.1073/pnas.77.10.5754 (1980).

821   29   Dhillon, S. & Keam, S. J. Bremelanotide: First Approval. *Drugs* **79**, 1599-1606,

822   doi:10.1007/s40265-019-01187-w (2019).

823   30   Ratner, M. Patients with porphyria bask in sunlight of FDA approval. *Nature biotechnology* **37**,

824   1390-1391, doi:10.1038/s41587-019-0347-0 (2019).

825   31   Liang, Y.-L. *et al.* Dominant Negative G Proteins Enhance Formation and Purification of Agonist-

826   GPCR-G Protein Complexes for Structure Determination. *ACS Pharmacology & Translational*

827   *Science* **1**, 12-20, doi:10.1021/acsptsci.8b00017 (2018).

828   32   Liang, Y.-L. *et al.* Phase-plate cryo-EM structure of a biased agonist-bound human GLP-1

829   receptor-Gs complex. *Nature* **555**, 121-125, doi:10.1038/nature25773 (2018).

830   33   Duan, J. *et al.* Cryo-EM structure of an activated VIP1 receptor-G protein complex revealed by

831   a NanoBiT tethering strategy. *Nature communications* **11**, 4121, doi:10.1038/s41467-020-

832   17933-8 (2020).

833   34   Rasmussen, S. G. F. *et al.* Crystal structure of the β2 adrenergic receptor-Gs protein complex.

834   *Nature* **477**, 549-555, doi:10.1038/nature10361 (2011).

835   35   Ballesteros, J. A. & Weinstein, H. in *Methods in Neurosciences* Vol. 25 (ed Stuart C. Sealfon)

836   366-428 (Academic Press, 1995).

837   36   Weis, W. I. & Kobilka, B. K. The Molecular Basis of G Protein-Coupled Receptor Activation.

838   *Annual review of biochemistry* **87**, 897-919, doi:10.1146/annurev-biochem-060614-033910

839   (2018).

840   37   Chrencik, J. E. *et al.* Crystal Structure of Antagonist Bound Human Lysophosphatidic Acid

841   Receptor 1. *Cell* **161**, 1633-1643, doi:10.1016/j.cell.2015.06.002 (2015).

842   38   Hua, T. *et al.* Crystal Structure of the Human Cannabinoid Receptor CB(1). *Cell* **167**, 750-

843   762.e714, doi:10.1016/j.cell.2016.10.004 (2016).

844   39   Hua, T. *et al.* Crystal structures of agonist-bound human cannabinoid receptor CB(1). *Nature*

845   **547**, 468-471, doi:10.1038/nature23272 (2017).

846   40   Krishna Kumar, K. *et al.* Structure of a Signaling Cannabinoid Receptor 1-G Protein Complex.

847   *Cell* **176**, 448-458.e412, doi:10.1016/j.cell.2018.11.040 (2019).

848   41   Hua, T. *et al.* Activation and Signaling Mechanism Revealed by Cannabinoid Receptor-G(i)

849   Complex Structures. *Cell* **180**, 655-665.e618, doi:10.1016/j.cell.2020.01.008 (2020).

850   42   Hanson, M. A. *et al.* Crystal structure of a lipid G protein-coupled receptor. *Science (New York,*

851   *N.Y.)* **335**, 851-855, doi:10.1126/science.1215904 (2012).

852   43   Yu, J. *et al.* Determination of the melanocortin-4 receptor structure identifies Ca(2+) as a

853   cofactor for ligand binding. *Science (New York, N.Y.)* **368**, 428-433,

27

854   doi:10.1126/science.aaz8995 (2020).

855   44   Flock, T. *et al.* Universal allosteric mechanism for Gα activation by GPCRs. *Nature* **524**, 173-179,
856   doi:10.1038/nature14663 (2015).

857   45   Yang, Y. K. *et al.* Molecular determinants of ligand binding to the human melanocortin-4
858   receptor. *Biochemistry* **39**, 14900-14911, doi:10.1021/bi001684q (2000).

859   46   Yang, Y. *et al.* Molecular determinants of human melanocortin-4 receptor responsible for
860   antagonist SHU9119 selective activity. *The Journal of biological chemistry* **277**, 20328-20335,
861   doi:10.1074/jbc.M201343200 (2002).

862   47   Zhou, Q. *et al.* Common activation mechanism of class A GPCRs. *eLife* **8**,
863   doi:10.7554/eLife.50279 (2019).

864   48   Rasmussen, S. G. *et al.* Structure of a nanobody-stabilized active state of the beta(2)
865   adrenoceptor. *Nature* **469**, 175-180, doi:10.1038/nature09648 (2011).

866   49   Garcia-Nafria, J., Lee, Y., Bai, X., Carpenter, B. & Tate, C. G. Cryo-EM structure of the adenosine
867   A2A receptor coupled to an engineered heterotrimeric G protein. *eLife* **7**,
868   doi:10.7554/eLife.35946 (2018).

869   50   Christopher, J. A. *et al.* Biophysical fragment screening of the β1-adrenergic receptor:
870   identification of high affinity arylpiperazine leads using structure-based drug design. *Journal of*
871   *medicinal chemistry* **56**, 3446-3455, doi:10.1021/jm400140q (2013).

872   51   Su, M. *et al.* Structural Basis of the Activation of Heterotrimeric Gs-Protein by Isoproterenol-
873   Bound β(1)-Adrenergic Receptor. *Molecular cell* **80**, 59-71.e54,
874   doi:10.1016/j.molcel.2020.08.001 (2020).

875   52   Ishchenko, A. *et al.* Toward G protein-coupled receptor structure-based drug design using X-
876   ray lasers. *IUCrJ* **6**, 1106-1119, doi:10.1107/s2052252519013137 (2019).

877   53   Ihara, K. *et al.* Isoprenoid-chained lipid EROCOC(17+4): a new matrix for membrane protein
878   crystallization and a crystal delivery medium in serial femtosecond crystallography. *Scientific*
879   *reports* **10**, 19305, doi:10.1038/s41598-020-76277-x (2020).

880   54   Nojima, S. *et al.* Cryo-EM Structure of the Prostaglandin E Receptor EP4 Coupled to G Protein.
881   *Structure (London, England : 1993)*, doi:10.1016/j.str.2020.11.007 (2020).

882   55   Toyoda, Y. *et al.* Ligand binding to human prostaglandin E receptor EP(4) at the lipid-bilayer
883   interface. *Nature chemical biology* **15**, 18-26, doi:10.1038/s41589-018-0131-3 (2019).

884   56   Yang, F. *et al.* Structural basis of GPBAR activation and bile acid recognition. *Nature*,
885   doi:10.1038/s41586-020-2569-1 (2020).

886   57   (!!! INVALID CITATION !!! (7, 51)).

887   58   Tao, Y. X. Molecular mechanisms of the neural melanocortin receptor dysfunction in severe
888   early onset obesity. *Molecular and cellular endocrinology* **239**, 1-14,
889   doi:10.1016/j.mce.2005.04.012 (2005).

890   59   Liang, Y. L. *et al.* Phase-plate cryo-EM structure of a biased agonist-bound human GLP-1
891   receptor-Gs complex. *Nature* **555**, 121-125, doi:10.1038/nature25773 (2018).

892   60   Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved
893   cryo-electron microscopy. *Nature methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

894   61   Zhang, K. Gctf: Real-time CTF determination and correction. *Journal of structural biology* **193**,
895   1-12, doi:10.1016/j.jsb.2015.11.003 (2016).

896   62   Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure
897   determination. *Journal of structural biology* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006

898   (2012).

899   63   Heymann, J. B. Guidelines for using Bsoft for high resolution reconstruction and validation of
900   biomolecular structures from electron micrographs. *Protein science : a publication of the*
901   *Protein Society* **27**, 159-171, doi:10.1002/pro.3293 (2018).

902   64   Zhao, L. H. *et al.* Structure and dynamics of the active human parathyroid hormone receptor-
903   1. *Science (New York, N.Y.)* **364**, 148-153, doi:10.1126/science.aav7942 (2019).

904   65   Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta*
905   *crystallographica. Section D, Biological crystallography* **60**, 2126-2132,
906   doi:10.1107/s0907444904019158 (2004).

907   66   Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular
908   structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221,
909   doi:10.1107/s0907444909052925 (2010).

910   67   Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis.
911   *Protein science : a publication of the Protein Society* **27**, 14-25, doi:10.1002/pro.3235 (2018).

912   68   Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and
913   developers. *Protein science : a publication of the Protein Society* **30**, 70-82,
914   doi:10.1002/pro.3943 (2021).

915   69   Lemmon, G. & Meiler, J. Rosetta Ligand docking with flexible XML protocols. *Methods in*
916   *molecular biology (Clifton, N.J.)* **819**, 143-155, doi:10.1007/978-1-61779-465-0_10 (2012).

917   70   Davis, I. W. & Baker, D. RosettaLigand docking with full ligand and receptor flexibility. *Journal*
918   *of molecular biology* **385**, 381-392, doi:10.1016/j.jmb.2008.11.010 (2009).

919   71   O'Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *Journal of cheminformatics* **3**, 33,
920   doi:10.1186/1758-2946-3-33 (2011).

921

933   **Author contributions** H.Z. designed the expression constructs, purified all the

934   MC4R−G$_s$ complexes, prepared the final samples for negative staining EM and cryo-

935 EM data collection, and participated in manuscript preparation; H.Z., D.D.S. and C.M.

936 performed specimen screening, cryo-EM grid preparation and data collection, map

937 calculation; H.Z., L.C., Q.S. and C.M. performed structural analysis and figures

938 preparation; H.Z., L.C., S.Y., W.F., A.D., S.X., Y. Zhou and D.Y. conducted mutation

939 studies, ligand binding and signaling experiments; Y.Z. and H.Z. analyzed the data with

940 the support of L.C., J.S., D.H.S., D.Y., H.E.X, T.H., T.Z. and M.-W.W.; Y.Z., M.-W.W.

941 and T.Z. supervised the project and wrote the manuscript with inputs from all co-

942 authors.

943 **Declaration of interests**

944 The authors declare no conflict of interest.
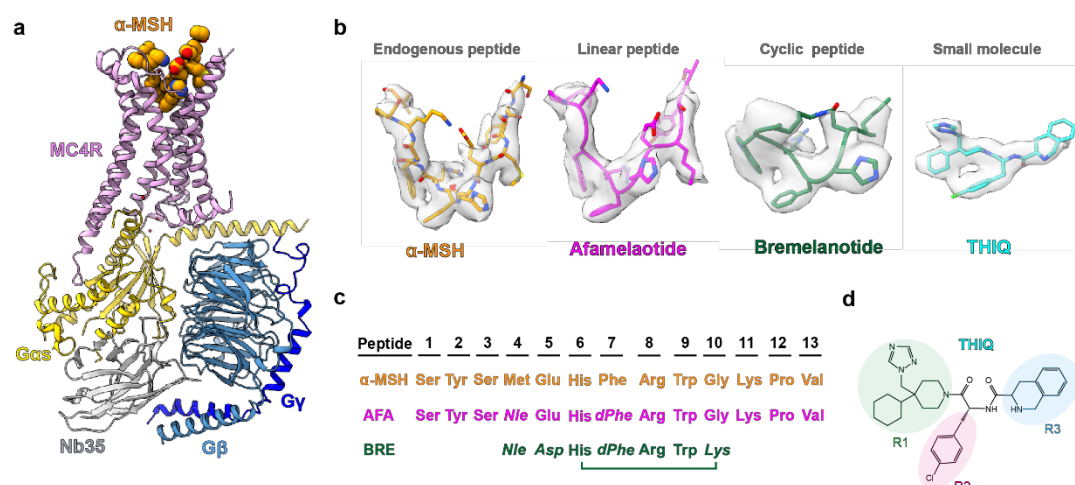
Fig. 1 Cryo-EM structures of the MC4R–G$_s$ complexes. a Cartoon representation of the α-MSH–bound MC4R–G$_s$ complex. MC4R, plum; α-MSH, orange; Gα$_s$, yellow; Gβ, cornflower blue; Gγ, blue; Nb35, grey. **b** Cryo-EM densities and models of α-MSH, peptide analogs (afamelanotide, bremelanotide) and small-molecule THIQ from their respective MC4R–G$_s$ complexes. α-MSH, orange; afamelanotide, magenta; bremelanotide, forest green; THIQ, cyan, density is shown at 0.025 contour level. **c** Amino acid sequence of the agonists used in this study. Sequence alignment of three peptide agonists of MC4R (left). The mutant residues of afamelanotide (AFA) and bremelanotide (BRE) relative to α-MSH are shown in italics (right). **d** Two-dimensional chemical structure of THIQ.

1
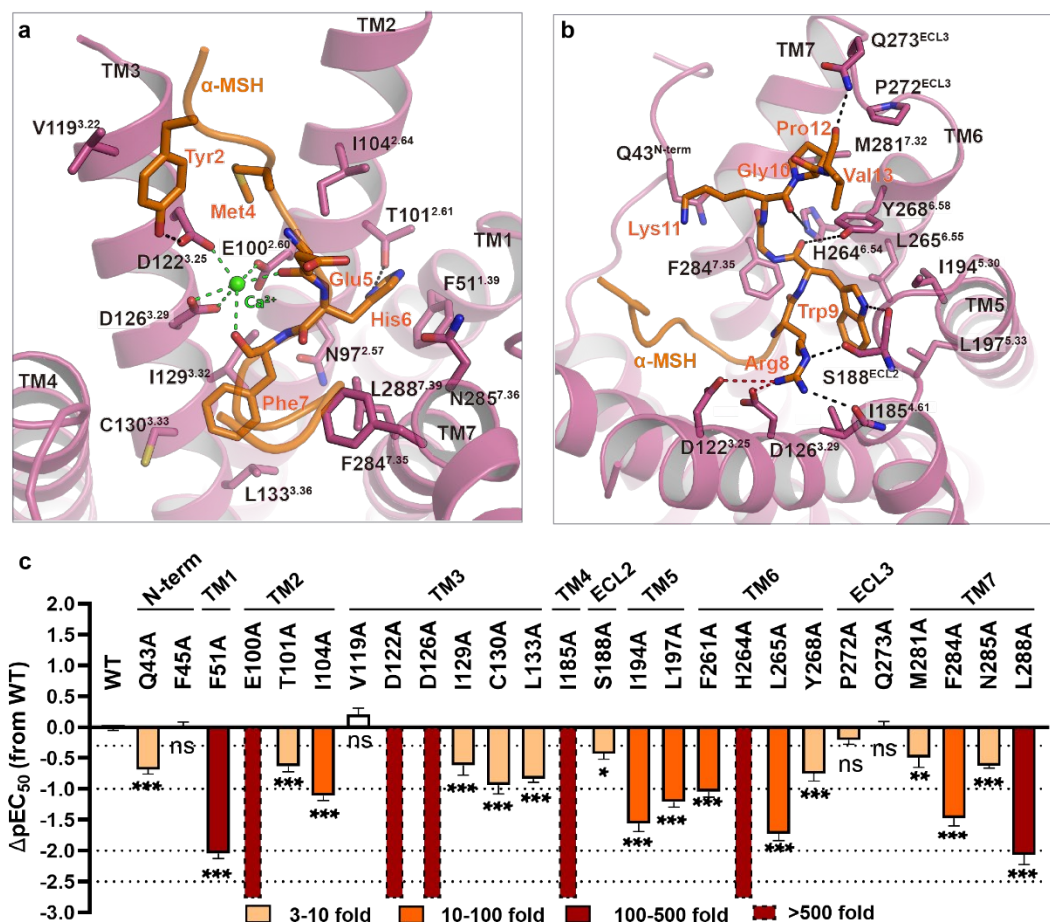
Fig. 2 The binding pocket of α-MSH in MC4R. a-b Detailed interactions of α-MSH (orange) with MC4R (plum). Structure viewed from the extracellular side shows the interaction network between MC4R, α-MSH, and Ca^{2+}. **c** α-MSH–induced cAMP accumulation assays of the residues involved in α-MSH binding. Bars represent differences in calculated α-MSH potency [pEC_{50}] for representative mutants relative to the wild-type receptor (WT). Data are colored according to the extent of effect. [ns]$P > 0.01$, *$P < 0.01$, **$P < 0.001$ and ***$P < 0.0001$ (one-way ANOVA followed by Dunnett's multiple comparisons test, compared with the response of WT). See Supplementary Table S4 for detailed statistical evaluation and receptor expression levels.
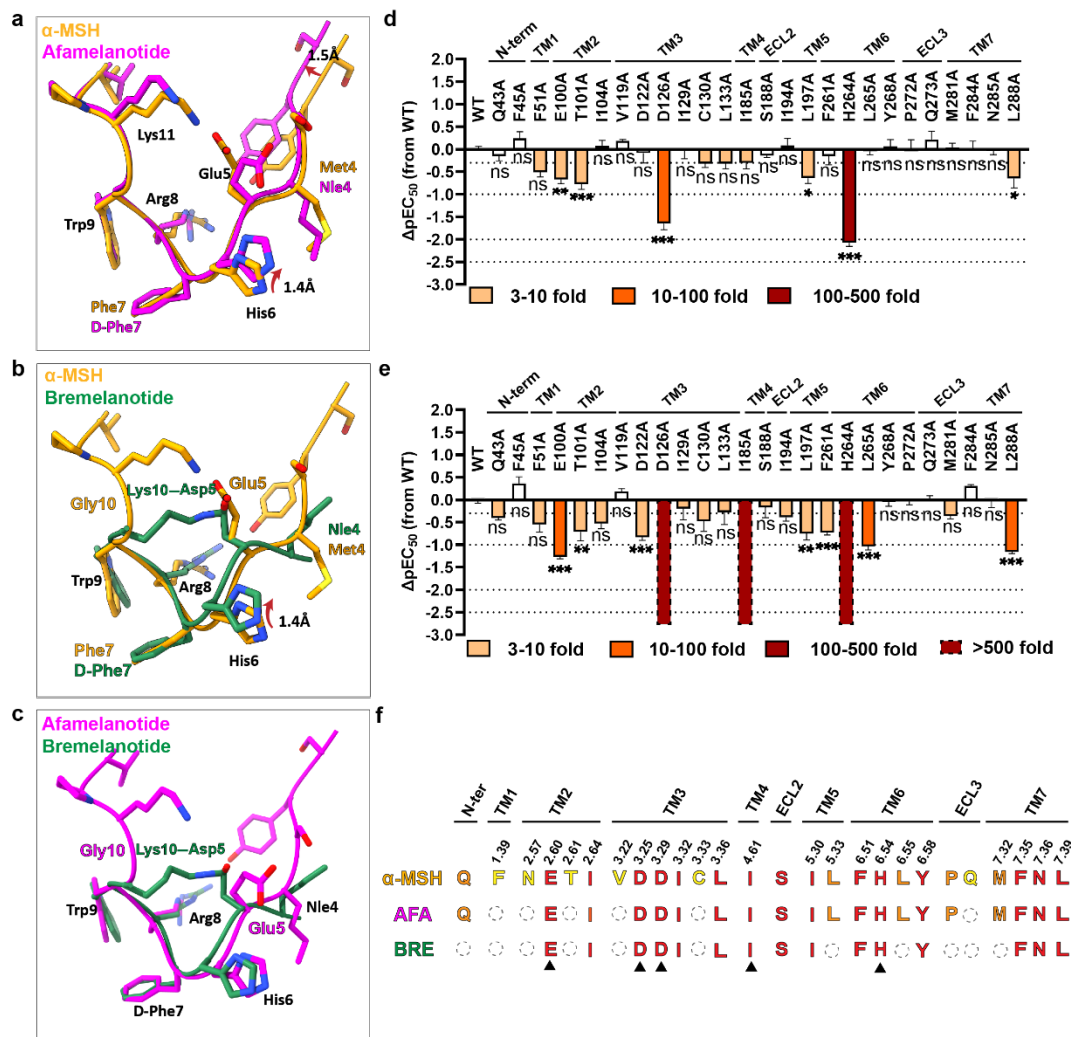
Fig. 3 Differences in the recognition mode of three peptide agonists of MC4R.

a-c Parallel structural comparisons of α-MSH–afamelanotide, α-MSH–bremelanotide and afamelanotide–bremelanotide. The peptide ligands are superimposed upon alignment of MC4R. α-MSH, orange; afamelanotide, magenta; bremelanotide, forest green. **d-e** Afamelanotide- (**d**) and bremelanotide-induced cAMP accumulation (**e**) assays of the residues involved in ligand recognition. Bars represent differences in calculated agonist potency [pEC_{50}] for each mutant relative to the wild-type receptor (WT). Data are colored according to the extent of effect. $^{ns}P > 0.01$, $*P < 0.01$, $**P < 0.001$ and $***P < 0.0001$ (one-way ANOVA followed by Dunnett's multiple comparisons test, compared with the response of WT). See Supplementary Table S7 and S8 for detailed statistical evaluation and receptor expression levels. **f** Comparison of the binding site of α-MSH, afamelanotide (AFA) and bremelanotide (BRE) in MC4R. The

triangle labeled residues are crucial for MC4R activation.
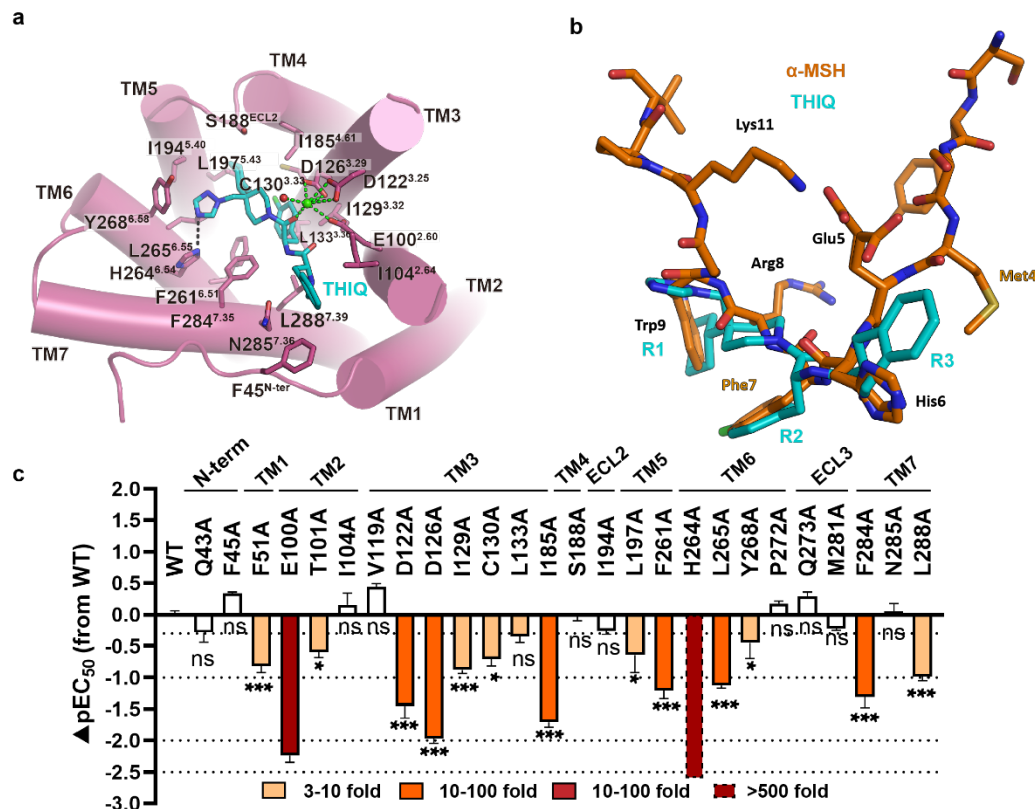
Fig. 4 Interaction between THIQ and MC4R. a Detailed interaction of small-molecule THIQ (cyan) with MC4R (plum). Structure viewed from the extracellular side shows the interaction network between MC4R, THIQ and Ca^{2+}. **b** Structure comparison of α-MSH with THIQ. The two ligands are aligned by the receptor. **c** THIQ-induced cAMP accumulation assays of the residues involved in THIQ binding. Bars represent differences in calculated α-MSH potency [pEC_{50}] for each mutant relative to the wild-type receptor (WT). Data are colored according to the extent of effect. [ns]$P > 0.01$, *$P <$ 0.01, **$P < 0.001$ and ***$P < 0.0001$ (one-way ANOVA followed by Dunnett's multiple comparisons test, compared with the response of WT). See Supplementary Table S10 and S11 for detailed statistical evaluation and receptor expression levels.
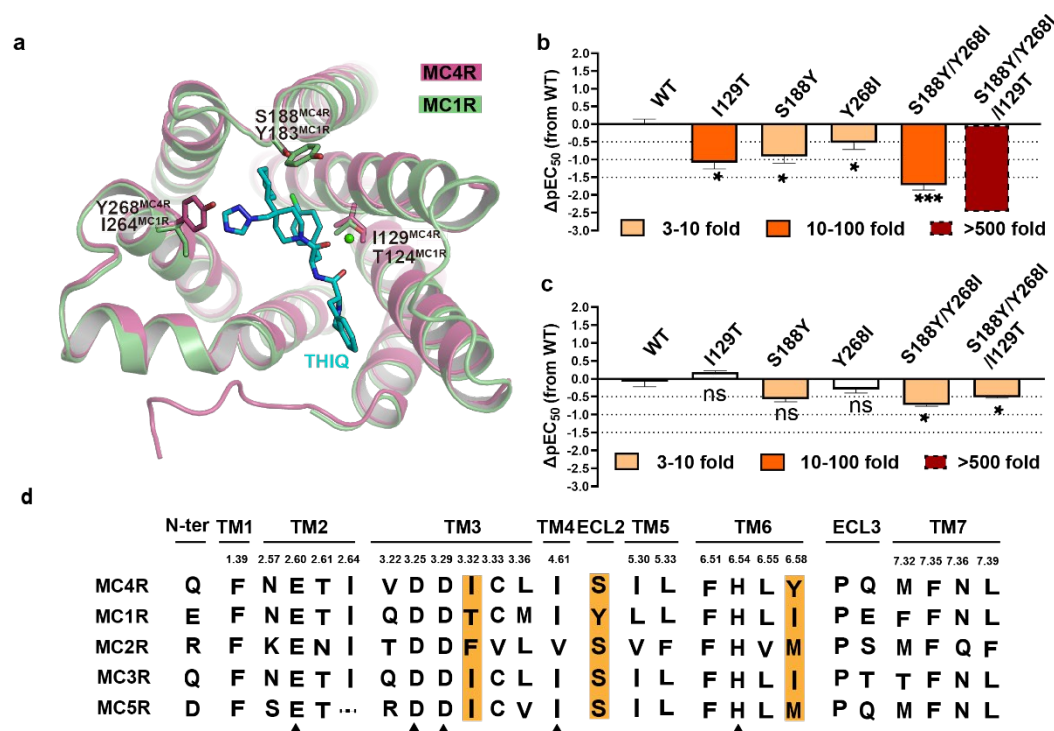
Fig. 5 Selective activation of MC4R by THIQ. a Structure comparison of MC4R and the homology modeled structure of MC1R. The non-conserved residues in the THIQ binding pockets between the two receptors are shown as sticks. MC4R, plum; MC1R, green; THIQ, cyan. **b-c** THIQ– (**b**) and α-MSH–induced cAMP accumulation (**c**) assays of the potential residues involved in selectivity. Bars represent differences in calculated α-MSH potency [pEC$_{50}$] for representative mutants relative to the wild-type receptor (WT). $^{ns}P > 0.01$, $*P < 0.01$, $**P < 0.001$ and $***P < 0.0001$ (one-way ANOVA followed by Dunnett's multiple comparisons test, compared with the response of WT). See Supplementary Table S4 and S11 for detailed statistical evaluation and receptor expression levels. **d** Sequence alignment of the residues in the orthostatic binding pocket of MC4R and other MCR family members. Highlighted are the orthostatic binding pocket residues that contribute to selectivity.

6
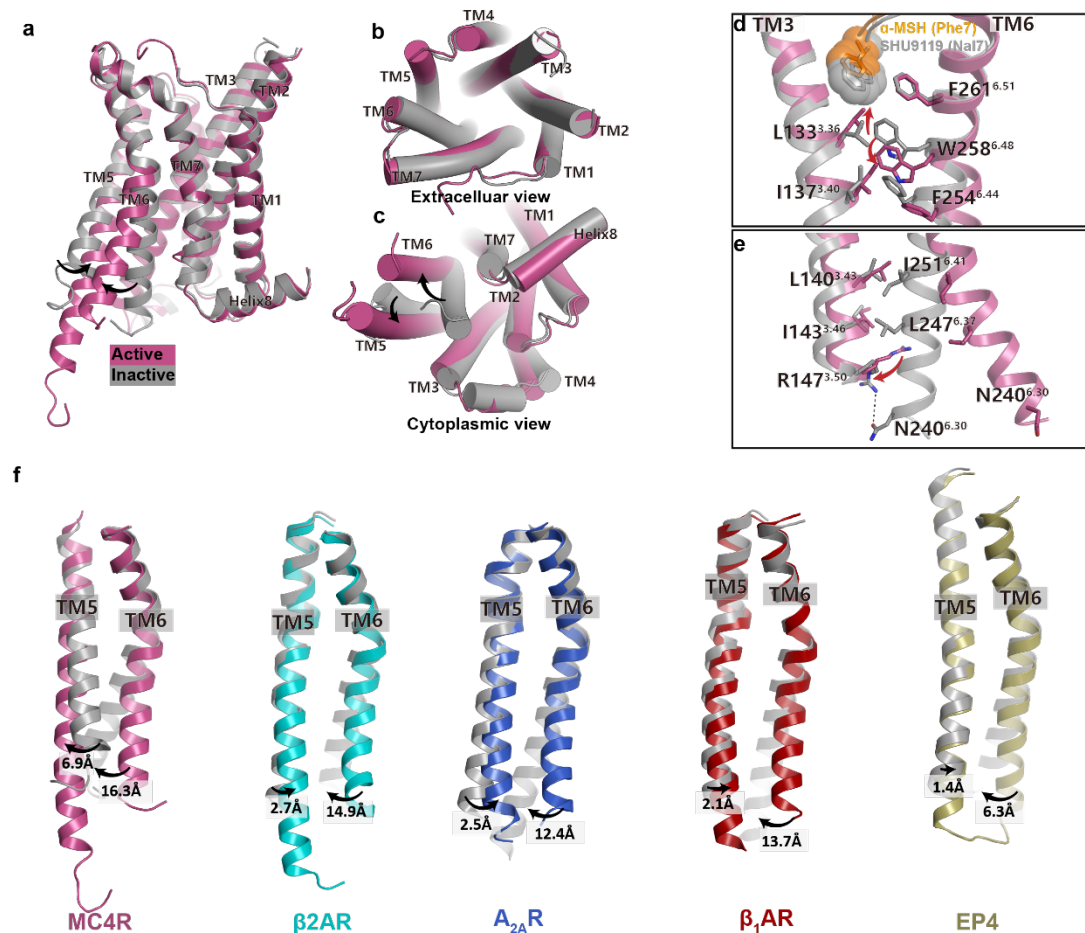
Fig. 6 Activation mechanism of MC4R. a-c Superposition of the α-MSH-activated MC4R (plum) with the antagonist (SHU9119) –bound MC4R (grey) (PDB: 6w25). (**a**) Side view. (**b**) Extracellular view. (**c**) Cytoplasmic view. **d-e** Close-up views of the conformational changes of the crucial residues involved in MC4R activation. **f** Structure comparisons of TM5 movements among peptide receptors during activation. Inactive and active MC4R, grey and plum; inactive and active β2AR, grey (PDB: 6PS6) and cyan (PDB: 3SN6); inactive and active A2AR, grey (PDB: 6LPJ) and blue (PDB: 6GDG); inactive and active β1AR, grey (PDB: 3ZPQ) and red (PDB: 7JJO); inactive and active EP4, grey (PDB: 5YWY) and yellow (PDB: 7D7M).
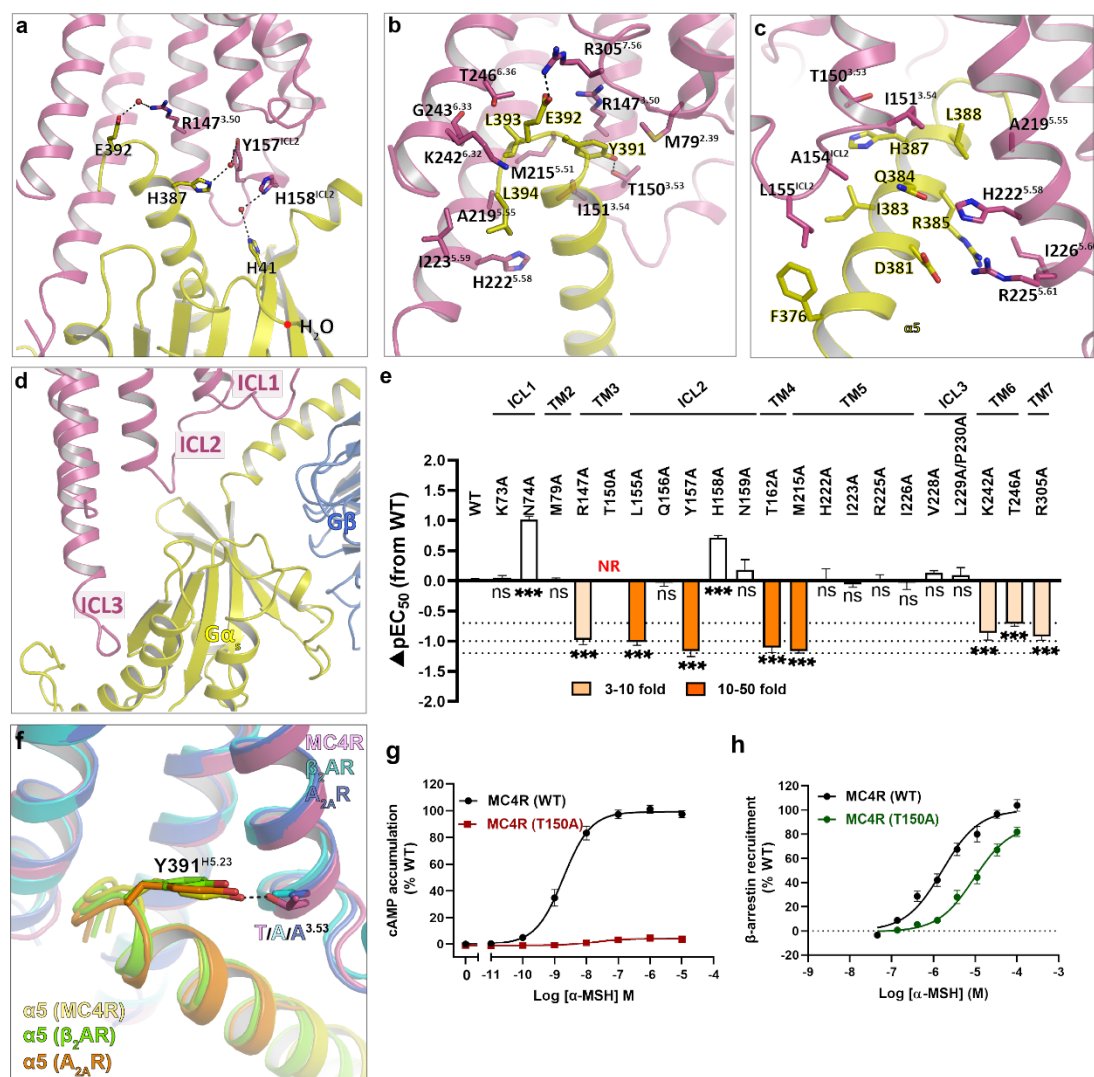
Fig. 7 MC4R-G_s protein coupling. a Water molecules in the MC4R-G_s interface. Hydrogen bonds are shown as dashed lines. **b-c** Detailed interaction between MC4R and G_s-α5 helix. **d** Detailed interaction between three intracellular loops of MC4R and Gαs. **e** α-MSH-induced cAMP accumulation assays of the G_s binding site in MC4R. Bars represent differences in calculated α-MSH potency [pEC_{50}] for each mutant relative to the wild-type receptor (WT). Data are colored according to the extent of effect. $^{ns}P > 0.01$, $*P < 0.01$, $**P < 0.001$ and $***P < 0.0001$ (one-way ANOVA followed by Dunnett's multiple comparisons test, compared with the response of WT). **f** $T^{3.53}$ of MC4R participates in the recognition of Y391 in G_s-α5 but not the corresponding $A^{3.53}$ of β2AR and A2AR. Plum cartoon, MC4R structure; cyan cartoon, β₂AR structure; blue cartoon, A₂AR structure. MC4R–G_s, plum (MC4R) and yellow (α5); β₂AR–G_s (PDB:

8

3SN6), cyan (β_2AR) and green (α5); $A_{2A}R$–G_s (PDB: 6GDG), blue ($A_{2A}R$) and orange (α5). **g** α-MSH–induced cAMP accumulation assays of WT and T150$^{3.53}$A mutant of MC4R. **h** α-MSH–induced β-arrestin2 recruitment assays of WT and T150$^{3.53}$A mutant of MC4R. See Supplementary Table S4 and S14 for detailed statistical evaluation and receptor expression levels.
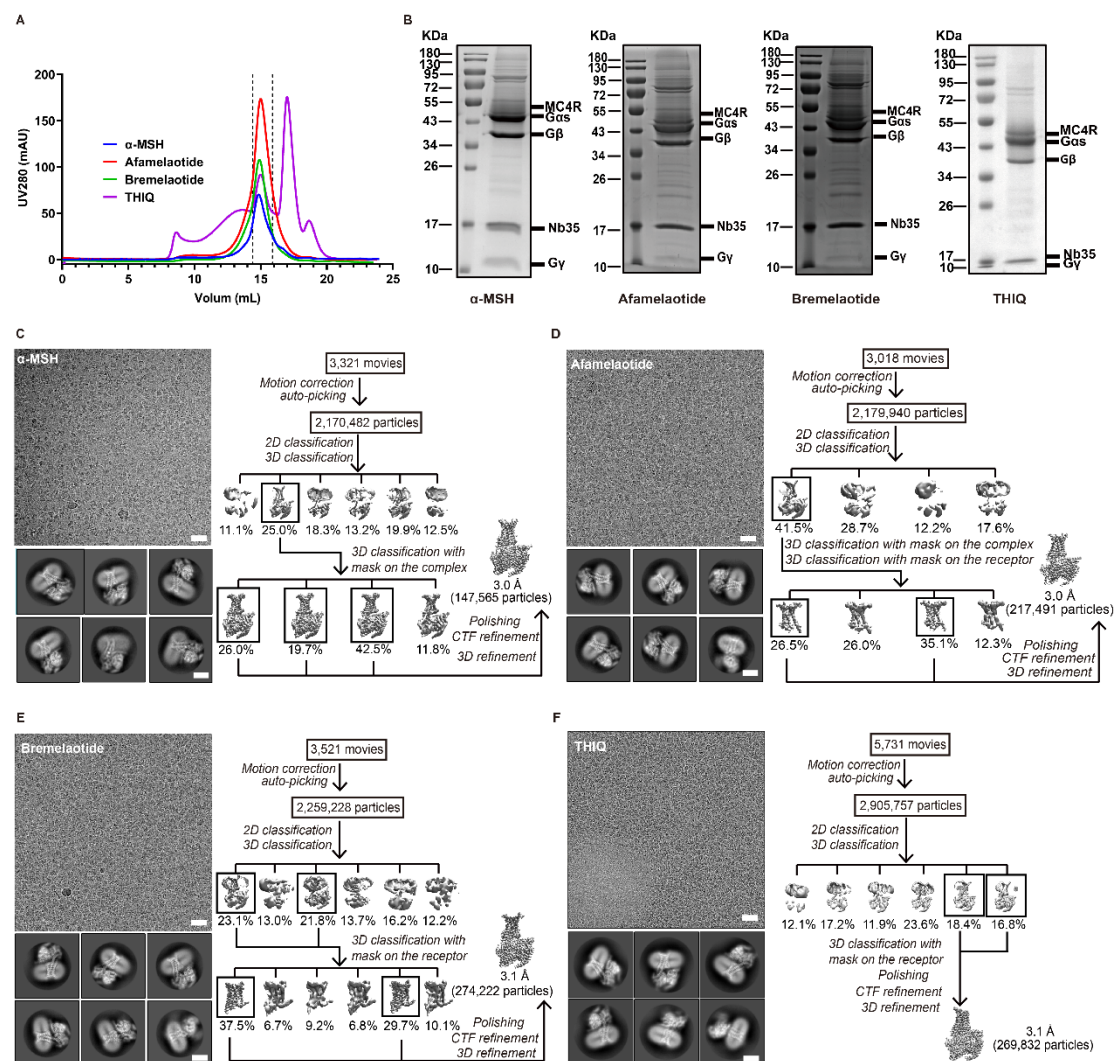
Supplementary information, Fig. S1| Purification and single-particle reconstruction of the agonists–bound MC4R–G$_s$ complexes.

a, Size exclusion chromatography (SEC) profile (a) and SDS-PAGE analysis (b) of α-MSH–, afamelanotide–, bremelanotide– and THIQ–bound MC4R–G$_s$ complexes. Fractions between two dashed lines in the SEC profile were pooled and concentrated for cryo-EM analysis. c-f Cryo-EM micrographs (scale bar: 30 nm), 2D class averages (scale bar: 5 nm) and the flow chart of cryo-EM data processing for α-MSH– (c), afamelanotide– (d), bremelanotide– (e) and THIQ–bound (f) MC4R–G$_s$ complexes.

10
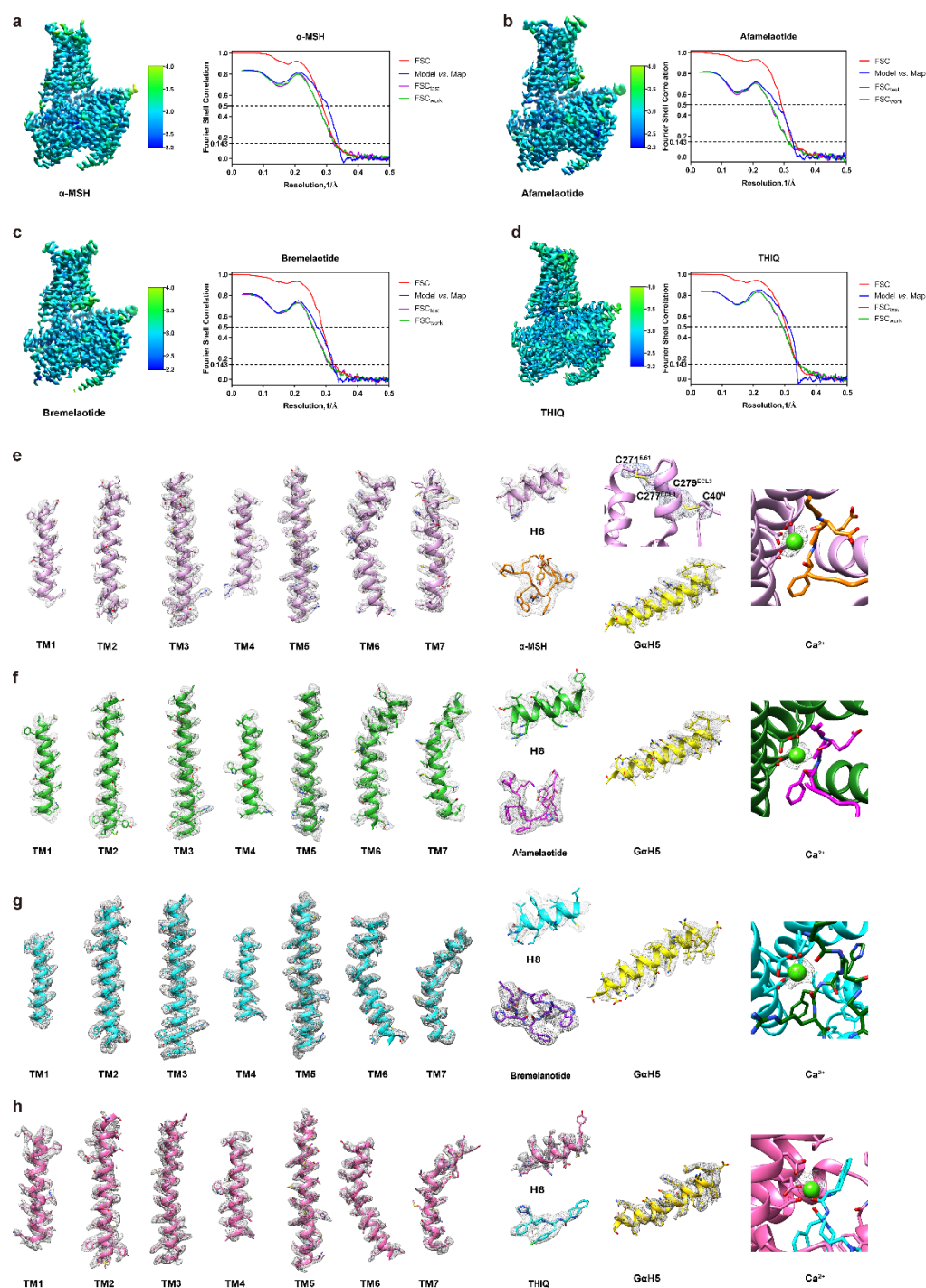
Supplementary information, Fig. S2| Overall resolution and cryo-EM density analysis of the MC4R-G$_s$ complexes.

a-d, Local resolution maps, gold-standard fourier shell correlation (FSC) curves, the model-vs-map, FSC$_{work}$ and FSC$_{test}$ validation curves of the α-MSH– (a), afamelanotide– (b), bremelanotide– (c) and THIQ–bound (d) MC4R–G$_s$ complexes.

e-h, EM density maps and models are shown for all seven-transmembrane helices,

11

helix 8, agonists, $G\alpha_s$ α5-helix and Ca^{2+} of the α-MSH– (e), afamelanotide– (f), bremelanotide– (g) and THIQ–bound (h) MC4R–G_s complexes. α-MSH–bound MC4R, plum; afamelanotide–bound MC4R, green; bremelanotide–bound MC4R, cyan; THIQ–bound MC4R, pink; α-MSH, orange; afamelanotide, magenta; bremelanotide, purple; THIQ, cyan; $G\alpha_s$-α5, yellow; Ca^{2+}, green. Densities are shown at 0.020-0.045 contour level, except the disulfide bond between C40 and C279 at 0.01 contour level.
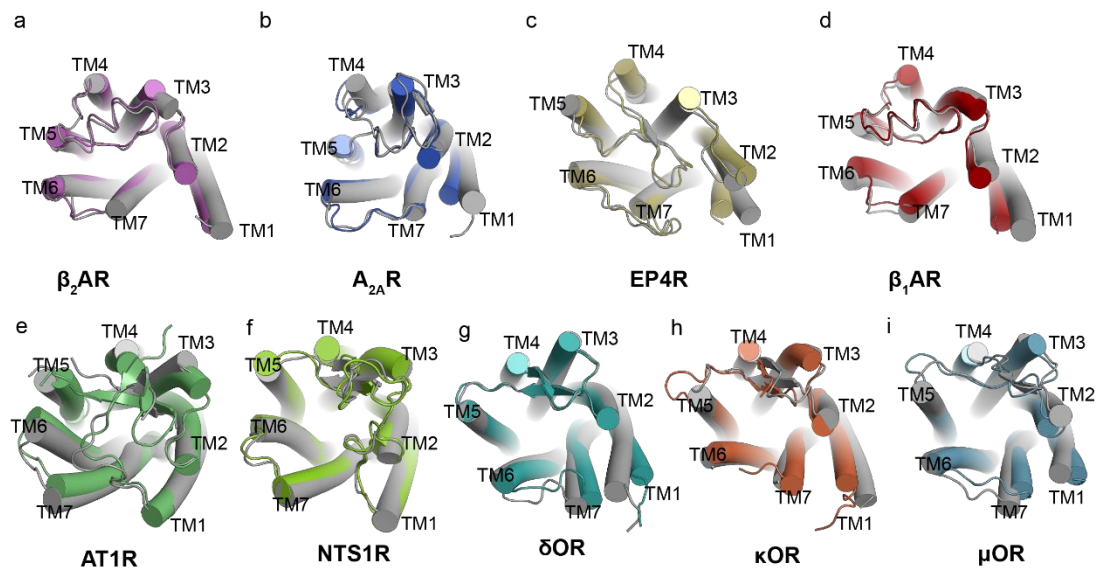
Supplementary information, Fig. S3| Structural comparison between active structures of class A G$_s$-coupled receptors and other peptide receptors.

a-d, Superimposition of inactive and active structures of class A G$_s$-coupled receptors. Inactive and active β$_2$AR, grey (PDB: 6PS6) and purple (PDB: 3SN6); inactive and active A$_{2A}$R, grey (PDB: 6LPJ) and blue (PDB: 6GDG); inactive and active EP4, grey (PDB: 5YWY) and yellow (PDB: 7D7M); inactive and active β$_1$AR, grey (PDB: 3ZPQ) and red (PDB: 7JJO). e-i, Superimposition of inactive and active structures of class A peptide receptors. Inactive and active AT1R, grey (PDB: 4YAY) and green (PDB: 6DO1); inactive and active NTSR1, grey (PDB: 4BUO) and yellow green (PDB: 6OS9); inactive and active δOR, grey (PDB: 4N6H) and turquoise (PDB: 6PT3); inactive and active κOR, grey (PDB: 4DJH) and orange (PDB: 6B73); inactive and active μOR, grey (PDB:4DKL) and steel blue (PDB:6DDE).

13
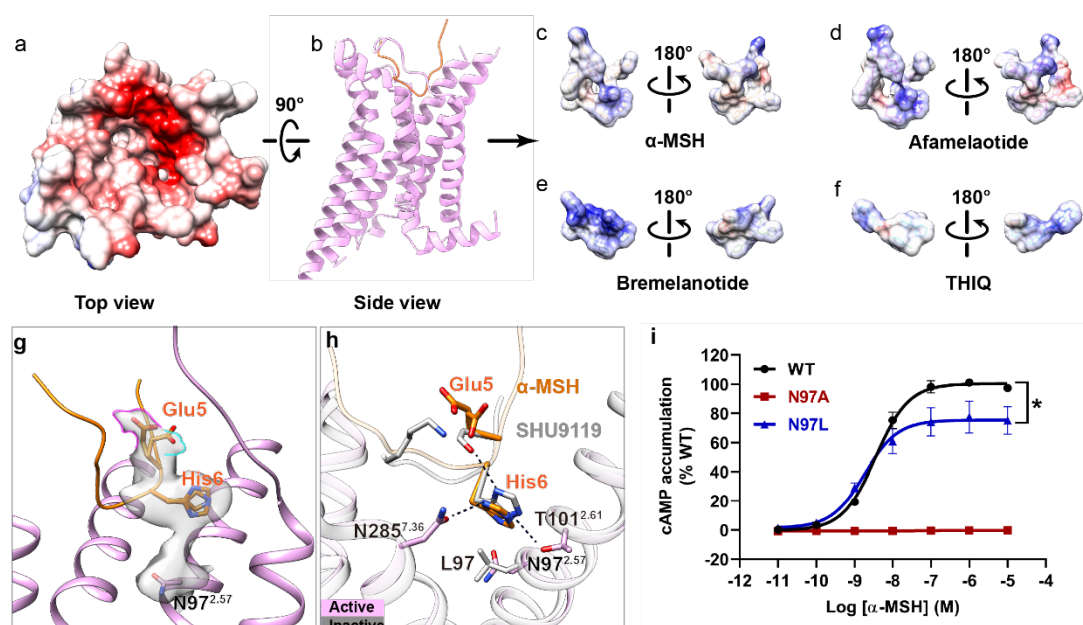
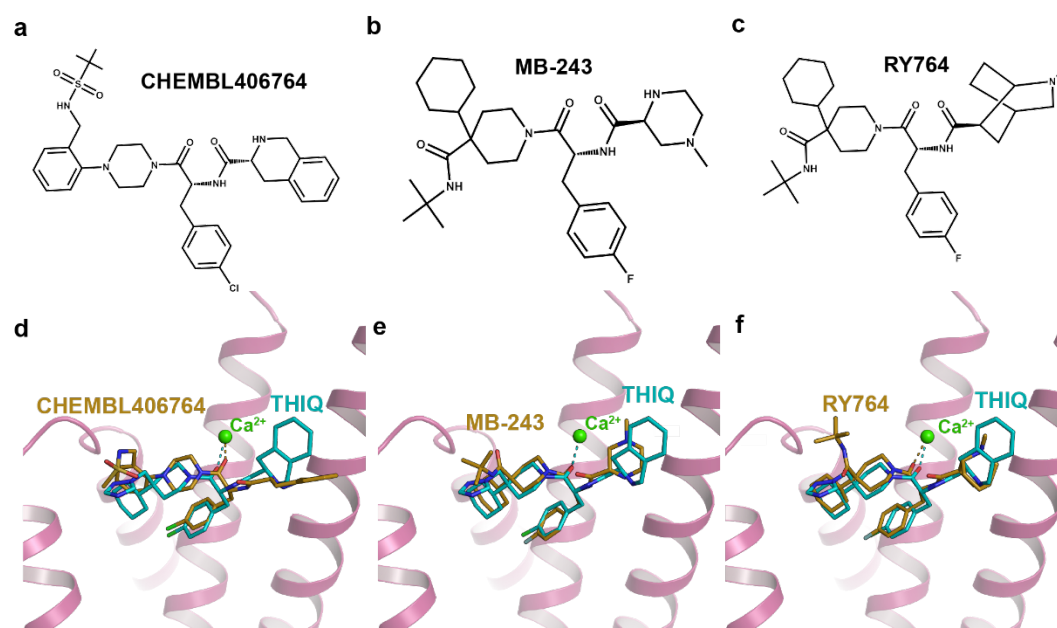**Supplementary information, Fig. S4| Ligand binding pocket of MC4R.**

a-f, The electrostatic potential analysis for the orthosteric pocket of MC4R (a-b) as well as α-MSH (c), afamelanotide (d), bremelanotide (e) and THIQ (f).

g, Cryo-EM densities of Glu5 and His6 in α-MSH and N97$^{2.57}$ in MC4R are shown at the 0.011 contour level. h, Superposition of the α-MSH (orange)–activated MC4R (plum) with the antagonist (SHU9119, orange)-bound and thermostabilized MC4R (N97$^{2.57}$L) mutant (grey, PDB 6W25). i, α-MSH–induced cAMP accumulation assays of WT, N97A and N97L MC4Rs, *$P < 0.05$. (i) See Supplementary Table S4 and S5 for detailed statistical evaluation and receptor expression levels.
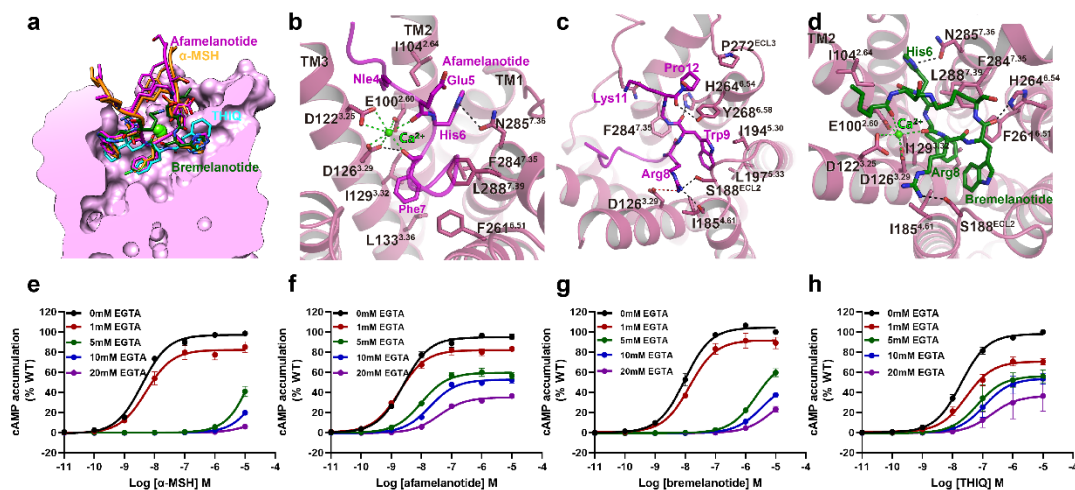
Supplementary information, Fig. S5| Binding sites of afamelanotide and bremelanotide in MC4R and effects of Ca^{2+} on agonist binding and agonist–induced signaling.

a, Comparative binding poses of four different ligands in the binding pocket. b-d, Detailed interaction between afamelanotide (b-c) and bremelanotide (d) with MC4R. MC4R, plum; afamelanotide, magenta; bremelanotide, sea green; Ca^{2+}, green. e-h, Effects of Ca^{2+} on α-MSH– (e), afamelanotide– (f), bremelanotide– (g) and THIQ– (h) induced cAMP accumulation.

15

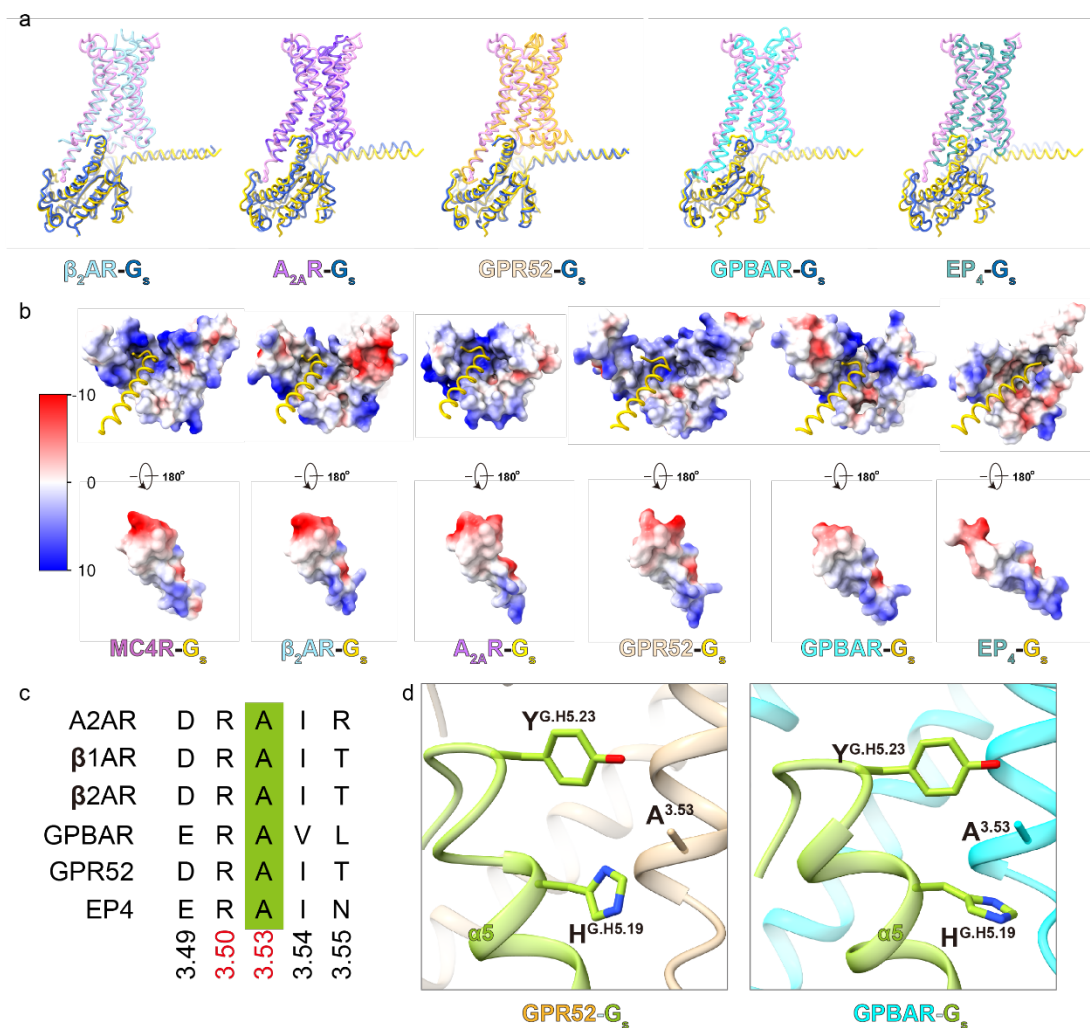**Supplementary information, Fig. S6| Molecular docking of MC4R agonists.**

a-c, Chemical structures of CHEMBL406764 (a), MB-243 (b) and RY764 (c);

d-f, Docking of CHEMBL406764 (d), MB-243 (e) and RY764 (f) in the THIQ–bound MC4R structure. The rank 1 binding pose of each ligand was selected to display.

Supplementary information, Fig. S7| Sequence alignment of the MCR family members and the orthosteric pocket difference between MC4R and MC1R.

a, Sequence alignment of the melanocortin family receptors. The sequence alignment was created using GPCRdb (http://www.gpcrdb. org). The residues of α-MSH binding

17

sites, G_s coupling sites and different residues between MC4R and MC1R in respective ligand-binding pocket are highlighted with hollow diamond, hollow triangle and solid triangle symbols. Secondary structure elements are annotated underneath the sequences based on the structure of MC4R. b, The orthosteric pocket of MC4R. The pocket surface is displayed in grey and the residues in the ligand pocket of MC4R that are different from the corresponding residues of MC1R are shown as balls (the orange balls represent the residues displaying different chemical properties from the corresponding residues of MC1R, the green balls represent residues with similar chemical properties). c-d, Differentiated residues in the orthosteric pocket between the MC4R structure and the homology model of MC1R. c, Highlighting view of the residues with different chemical properties. d, Highlighting view of the residues with similar chemical properties.
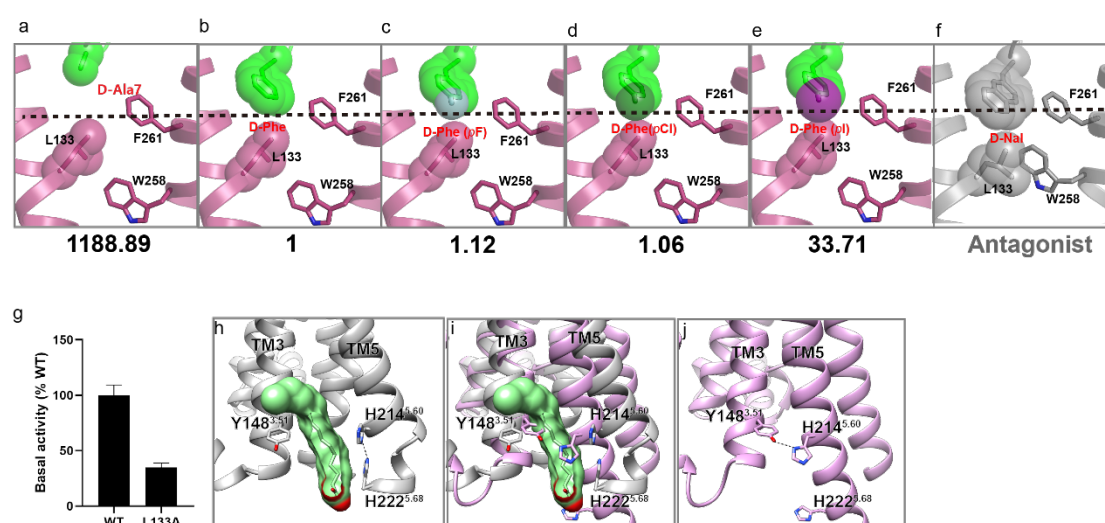
Supplementary information, Fig. S8| Relationships between chemical modification at the position of D-Phe residue and MC4R activation, and lipid in the inactive MC4R. a-f, Chemical replacement of D-Phe to D-Ala (a), D-Phe(pF) (c), D-Phe(pCl) (d) and D-Phe(pI) (e) based on the active structure of bremelanotide-bound MC4R (b). The D-Nal modification is exhibited using the inactive structure of SHU9119-MC4R (f). The fold change in EC_{50} value of each ligand compared to [Nle4, D-Phe7]α-MSH-NH$_2$ (b) is shown under the corresponding panel. Active MC4R, plum; inactive MC4R, grey. g, Basal activity of MC4R (wild type, WT) and MC4R (L133$^{3.36}$A). h, Lipid between TM3 and TM5 in the inactive MC4R. i, Superimposition of inactive and active structures of MC4R. j, Hydrogen bond between Y148$^{3.51}$ and H214$^{5.60}$ in the inactive MC4R.

Supplementary information, Fig. S9| Interaction patterns between class A GPCRs and Gα$_s$.

a, Pairwise comparison of MC4R–Gα$_s$ and other class A GPCR–Gα$_s$ structures. MC4R, plum; MC4R–bound Gα$_s$, yellow; other receptor-bound Gα$_s$, blue; β$_2$AR (PDB: 3SN6); sky blue; A$_{2A}$R (PDB: 6GDG), purple; GPR52 (PDB: 6LI3); orange, GPBAR (PDB: 7CFM); cyan, EP$_4$ (PDB: 7D7M), teal. b, The electrostatic potential analysis of the G protein interface for class A GPCR–G$_s$ complexes. c, Residues on TM3 involved in G protein recognition among for G$_s$ coupled class A receptors. The presentation was created using GPCRdb (http://www.gpcrdb.org). The conserved interaction residues of R$^{3.50}$ and A$^{3.53}$ are highlighted. d, Detailed interactions of receptor A$^{3.53}$ with the α5 helix of G protein in the structures of GPR52–G$_s$ and GPBAR–G$_s$.