Photoaffinity Cross-Linking and Unnatural Amino Acid Mutagenesis Reveal Insights into Calcitonin Gene-Related Peptide Binding to the Calcitonin Receptor-like Receptor/Receptor Activity-Modifying Protein 1 (CLR/RAMP1) Complex

John Simms,†‡ Romeo Uddin,† Thomas P. Sakmar,§ Joseph J. Gingell,|| Michael L. Garelja,|| Debbie L. Hay,‡∥ Margaret A. Brimble,‡∥ Paul W. Harris,|| Christopher A. Reynolds,⊥ and David R. Poyner‡∥

†Aston University, Birmingham B4 7ET, U.K.
‡Coventry University, Priory Street, Coventry CV1 5FB, U.K.
§The Rockefeller University, 1230 York Avenue, New York, New York 10065, United States
∥University of Auckland, 3A Symonds Street, Auckland 1010, New Zealand
⊥University of Essex, Wivenhoe Park, Colchester CO4 3SQ, U.K.

ABSTRACT: Calcitonin gene-related peptide (CGRP) binds to the complex of the calcitonin receptor-like receptor (CLR) with receptor activity-modifying protein 1 (RAMP1). How CGRP interacts with the transmembrane domain (including the extracellular loops) of this family B receptor remains unclear. In this study, a photoaffinity cross-linker, p-azido l-phenylalanine (azF), was incorporated into CLR, chiefly in the second extracellular loop (ECL2) using genetic code expansion and unnatural amino acid mutagenesis. The method was optimized to ensure efficient photolysis of azF residues near the transmembrane bundle of the receptor. A CGRP analogue modified with fluorescein at position 15 was used for detection of ultraviolet-induced cross-linking. The methodology was verified by confirming the known contacts of CGRP to the extracellular domain of CLR. Within ECL2, the chief contacts were I284 on the loop itself and L291, at the top of the fifth transmembrane helix (TM5). Minor contacts were noted along the lip of ECL2 between S286 and L290 and also with M223 in TM3 and F349 in TM6. Full length molecular models of the bound receptor complex suggest that CGRP sits at the top of the TM bundle, with Thr6 of the peptide making contacts with L291 and H295. I284 is likely to contact Leu12 and Ala13 of CGRP, and Leu16 of CGRP is at the ECL/extracellular domain boundary of CLR. The reduced potency, Emax, and affinity of [Leu16Ala]-human α CGRP are consistent with this model. Contacts between Thr6 of CGRP and H295 may be particularly important for receptor activation.

Calcitonin gene-related peptide (CGRP) is an abundant sensory neuropeptide; it is known to be a very potent vasodilator and is implicated in migraines. It is part of the calcitonin superfamily of peptides that also includes adrenomedullin, adrenomedullin 2/intermedin, and amylin.1,2 The peptides all bind to family B/secretin-like G protein-coupled receptors, either the calcitonin receptor or the calcitonin receptor-like receptor (CLR). However, for high-affinity binding for all of these ligands apart from calcitonin, it is necessary for the receptor to form a complex with a second transmembrane protein from the receptor activity-modifying protein (RAMP) family with only three members. The CGRP receptor is formed from CLR and RAMP1, although CGRP will also bind with high affinity to the complex of the calcitonin receptor and RAMP1.1

The binding of endogenous peptide ligands to family B GPCRs generally follows the two-domain model in which the N-terminal tail of the peptide interacts with the transmembrane (TM) domain of the receptor whereas the C-terminal tail binds to the extracellular domain (ECD) of the receptor.3 Recently, a number of structures of full length family B GPCRs have been published, namely, for the glucagon and glucagon-like peptide 1 (GLP-1) receptors.4–8 The full length calcitonin receptor has been visualized by cryo-electron microscopy, but the ECD of the receptor was very flexible and so could not be resolved. The bound salmon calcitonin peptide was also poorly resolved, and its binding mode was refined by only rigid body fitting of the calcitonin receptor ECD structure [Protein Data Bank (PDB) entry 5I10] and confirmed with data from mutagenesis.9 For the CGRP

Received: May 2, 2018
Revised: July 9, 2018
Published: July 13, 2018
receptor, a crystal structure is available showing the C-terminal tail of a CGRP analogue bound to the ECDs of CLR and RAMP1.10 Extensive mutagenesis data that have allowed the construction of models of how CGRP interacts with the TM domain of the receptor are available.11–13 These studies have implicated extracellular loop 2 (ECL2) of CLR as being especially important for CGRP binding.12–14 However, direct evidence of molecular contacts between the ligand and receptor is lacking. Photoaffinity cross-linking has long been used to map the binding of ligands to receptors.16 Conventionally, a photoaffinity cross-linker is incorporated into the ligand; this then requires extensive peptide mapping to identify the most likely contact point on the receptor. With the use of genetic code expansion and unnatural amino acid mutagenesis,17 it is now possible to incorporate the photoaffinity cross-linker into the receptor, thereby eliminating the need for peptide mapping or sequencing to identify contact points on the receptor. This general approach, known as targeted photo-cross-linking,18 has been used to map ligand–receptor interactions in a number of GPCRs, including the neurokinin NK-1 receptor,19 ghrelin receptor,20 type 1 receptor for the corticotropin releasing factor (CRF1R),21 the CGRP receptor. In the study presented here, we have developed a modified targeted photo-cross-linking strategy to study ECL2 of CLR and have identified two major contact points for CGRP: I284 and L291.

**MATERIALS AND METHODS**

**Materials.** Human α CGRP was from Bachem (St. Helens, U.K.). CGRP analogues were synthesized as indicated below. Alpha Screen, LANCE CAMP assay kits, reagents, plates, and 125I-labeled human iodohistidyl 10-ethyl-fluorescein.27 The crude peptide was cleaved from the resin with concomitant removal of side chain protecting groups by sample displacement chromatography28 on a C18 Gemini (Phenomenex) column (10 mm × 250 mm) at a flow rate of 5 mL/min at 50 °C to give 25.1 mg of pure linear (unoxidized) [Lys(S(6)-carboxyfluorescein)]15-α-CGRP.

The linear peptide (25 mg) was dissolved in 0.1 M Tris-HCl (pH 7.85) at a concentration of 1 mg/mL and the oxidation (disulfide formation) was allowed to proceed at room temperature open to air. Monitoring by reverse phase high-performance liquid chromatography (HPLC) indicated that the reaction was complete after 5 h. The reaction mixture was acidified to pH 2, diluted with water (60 mL), and purified by semipreparative reverse phase HPLC using a C18 Gemini column (10 mm × 250 mm) at a flow rate of 5 mL/min at 50 °C and eluted using an appropriate gradient based on the analytical HPLC profile. Fractions containing the pure peptide were identified by electrospray mass spectrometry and/or HPLC, pooled, and lyophilized to give [Lys(S(6)-carboxyfluorescein)]15-α-CGRP at >95% purity as judged by integration of the HPLC chromatogram at 210 nm (calculation mass of 4163.31 Da, found mass of 4162.20 Da).

**Unnatural Amino Acid Mutagenesis and Cross-Linking Strategy.** Unnatural amino acid mutagenesis was performed in HEK293T cells as described previously,24,25 using Amber mutant human CLR, myc-tagged RAMP1, suppressor tRNA, and azF aminocyl tRNA synthetase but with some modifications. HEK293T cells were seeded to a density of 80000 cells per well (resulting in 70% cell confluence the next day) of a six-well plate. Twenty-four hours after being seeded, the cells were transfected with a DNA cocktail comprising 1 μg of mutated/WT receptor, 1 μg of suppressor tRNA, and 0.5 μg of aminocyl tRNA synthetase using lipofectamine (Life Technologies catalog no. 18324-012) as per the manufacturer’s instructions. Four hours post-transfection, the medium was replaced with Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 20% (v/v) FBS and 0.5 mM azF. Cells were harvested after a further 24 h, and the membrane fraction was isolated by centrifugation. The membranes were incubated with the labeled peptide for 1 h at room temperature. Once the incubation was complete, the sample was split into two equal parts. One part was pipetted into a quartz cuvette and exposed to 254 nm light for 1 h (an exposure longer than that previously reported26), while the other was wrapped in aluminum foil and also placed in the light source. Heat was dissipated in both samples using silicone tubing inserted into the cuvette that was flushed with ice-cold water. Cross-linking was performed in a UVP cross-linker (Analytik Jena AG) at an energy of 100 μJ/cm² with the sample approximately 10 cm from the light source. Both the cross-linked samples and the control samples were pelleted by centrifugation, washed using phosphate-buffered saline, and solubilized using dodecyl maltoside for 4 h. Post-solubilization, the samples were centrifuged at 30000g for 30 min, after which 150 μL of the soluble fraction was pipetted into a well of a 96-well plate and read using a Mithras LB 940 instrument (Berthold). The cross-linked and control samples were compared to determine the degree of peptide labeling.
Tyrosine Mutagenesis. Individual amino acids in human HA-tagged CLR were mutated to tyrosine as described previously.\(^{11,12,14}\)

CAMP Production. Human HA-tagged CLR (WT or mutant) and myc-tagged RAMP1 were transfected into HEK293T or HEK293S cells, stimulated with human \(\alpha\)-CGRP or synthetic peptides, and CAMP accumulation was measured via Alpha Screen or LANCE as previously described.\(^{11,15,30}\)

Radioligand Binding. Competition radioligand binding assays were performed on HEK293S membranes using \(^{125}\)I-labeled human \(\alpha\)-CGRP. HEK293S cells were grown in 15 cm dishes, and each dish was transfected with 30 \(\mu\)g of human HA-tagged CLR and 30 \(\mu\)g of myc-tagged RAMP1 using polyethyleneimine (PEI) as described previously.\(^{13}\) Forty-eight hours after transfection, membranes were isolated by homogenizing cells followed by centrifugation at 100000 \(\times\) g for 1 h at 4 °C. Membranes were resuspended and stored at −80 °C. The binding assay was performed essentially as described previously except the membranes were incubated with labeled and unlabeled peptides for 1 h at room temperature.

Data Analysis. Concentration response curves were fitted to a three-parameter logistic equation to obtain \(E_{\text{max}}\), pEC\textsubscript{50}, and basal responses. The Hill coefficient was constrained to unity. Curve fitting was performed with GraphPad Prism version 7, as described previously.\(^{13}\) For photoaffinity cross-linking, paired Student’s t tests were used to assess if the total cross-linking for each mutant was statistically different from the nonspecific cross-linking, accepting significance at a \(P < 0.05\) level.

Molecular Modeling. Homology models of the active CLR were generated using Modeler\(^{32}\) utilizing the recently determined cryo-electron microscope structures of the activated GLP-1 and calcitonin receptors (PDB entries 5VAI and SUZ7, respectively).\(^{8,9}\) One thousand models combining homology-based restraints from both templates (>2 Å root-mean-square deviation between models) were generated by Modeler using the default method. The transmembrane and loop regions of the Modeler-based structures were further refined and ranked using the membrane relax module of ROSETTA.\(^{33}\) The refined models were clustered, and the best model was chosen on the basis of the largest cluster with the lowest ROSETTA membrane score.

The peptide was initially \(ab\ initio\) folded using an in-house script with a disulfide bond constraining Cys\(^2\) and Cys\(^5\) of the peptide. Regions of secondary structure\(^{34}\) were also imposed on the \(ab\ initio\) simulation. The best scoring folded peptide was docked on the basis of the FlexPepDock module of ROSETTA; 10000 docking solutions between the peptide and receptor were generated. The FlexPepDock docked solutions were filtered in two stages. The docked solutions were initially filtered using analogous contact restraints (cutoff of 5 Å) as observed in the 4RWG crystal structure.\(^{9,9}\) These results were further filtered by keeping any solution in which both Ile\(^{284}\) and Leu\(^{9}\) were within 5 Å of any atom of the docked CGRP ligand. The filtered results (737 docked poses) were then further refined using the membrane relax module also found in the ROSETTA suite of software. The middle structure of the largest, lowest-energy cluster was chosen for further analysis.

Replica exchange Monte Carlo simulations as implemented in Hippo\(^{35}\) were used explore the conformational space of the ROSETTA-refined model. Sixteen replicas were simulated over the temperature range of 300−500 K following a Boltzmann distribution in a 30 Å implicit membrane. Other variables were kept as the default settings. The trajectory was visualized using Visual Molecular Dynamics.\(^{6}\)

RESULTS

Activity of [Lys(5(6)-carboxyfluorescein)]\(^{15}\) Human \(\alpha\)-CGRP (15-Fluo CGRP). To identify cross-linking to CLR, a fluorescent CGRP agonist was required. On the basis of previously published structure−activity data,\(^{37}\) supported by molecular modeling,\(^{11,38}\) we reasoned that it should be possible to modify CGRP at position 15 while retaining agonist activity. The native Leu at this residue was replaced by Lys and derivatized with S(6)-carboxyfluorescein. The resulting analogue was evaluated for its ability to stimulate CAMP production via the CGRP receptor. While it had reduced potency, it was a full agonist (Figure 1), suggesting that this modification did not greatly affect the functional properties of the ligand.

Figure 1. CAMP production on HEK293T cells transfected with CLR and RAMP1 and exposed to CGRP or 15-Fluo CGRP for 15 min. Values are means ± the standard error of the mean (SEM); \(n = 3\) independent experiments.

Confirmation of the Cross-Linking Strategy. Preliminary experiments indicated that optimal specific cross-linking was obtained following a 1 h exposure to ultraviolet (UV) light. The availability of a crystal structure showing the binding of a modified CGRP analogue to the extracellular domain of the CGRP receptor\(^{10}\) provided a system for evaluating our methodology. A series of azF mutants was constructed in which residues within 5 Å of or more than 7 Å from the bound CGRP analogue were substituted. There was excellent agreement between the ability of the mutants to cross-link to CGRP and their distances from the bound CGRP analogue in the crystal structure (Figure 2).

Cross-Linking by Residues in ECL2. A scan was then performed on ECL2 of CLR and its flanking regions from H270 to Y292, excluding C282, which is involved in a conserved disulfide bond with C212 at the extracellular end of TM3. The data suggest that major points of attachment for 15-Fluo CGRP are I284 and L291, with much weaker although significant labeling in a number of other positions at the distal end of the loop (Figure 3). There was virtually no specific labeling at any position in the loop proximal to D280. It is possible that a lack of labeling might simply reflect the fact that the receptor is nonfunctional when substituted with azF. To test this, a Tyr scan of the residues mutated to azF was performed, excluding the native tyrosines at positions 277 and 278 (Figure 4). A Tyr rather than an azF scan was performed.
because the modified aminoacyl tRNA synthetase used to incorporate azF will also accept Tyr, albeit at lower efficiency, resulting in a mixed population of receptors potentially with one of two different residues at the same site. Tyr is structurally the closest natural amino acid to azF. The results showed that all the mutants assessed should still be able to interact with 15-Fluo CGRP at 100 nM, the concentration used in our cross-linking experiments, as the Tyr mutants retained a good sensitivity to CGRP.

**Role of TM3 and TM6.** The current cross-linking data and previous mutation studies have indicated that ECL2 and the first turn of TM5 are important in forming the binding site for CGRP. We extended the use of our cross-linking method to address the role of other parts of the receptor. M223 (TM3) and F349 (TM6) were mutated to azF. These residues are likely to face into the TM bundle of CLR but are approximately two helical turns deeper into the TM bundle than L291 is (Figure 6). Thus, they provide insight into how far CGRP penetrates into the TM core of CLR. The specific cross-linking seen with both of these residues was comparable with that of S286 (Figure 5), which appears to act as a minor contact in ECL2 (Figure 3), although we cannot exclude a lower cross-linking efficiency due to poor photolysis of residues deep in the TM bundle. It seems likely that M223 and F349 are at the base of the binding pocket for CGRP.

**Molecular Modeling of Bound CGRP.** Our modeling suggests that CGRP extends down as far as H295, in TM5. I284 is close to Leu12 and Ala13 of CGRP, and L291 is adjacent to Thr6 of CGRP (Figure 6). Leu16 of CGRP is the most proximal residue of the bound CGRP to make potential contact with the ECLs. It is close to I284 and V198 (in ECL1) but is also in the proximity of I32 of the ECD, perhaps making contact with both ECLs and the ECD of CLR.

**Activity of [Leu12Ala]- and [Leu16Ala]-Human α-CGRP.** To further test our proposed mode of binding, we investigated the ability of [Leu12Ala]- and [Leu16Ala]-human α-CGRP to stimulate cAMP. Both analogues showed substantial reductions in potency and $E_{\text{max}}$ with [Leu16Ala]-human α-CGRP showing the largest effects (Figure 7). This result was confirmed in a radioligand binding assay, in which [Leu16Ala]-human α-CGRP also showed a substantial reduction in affinity.
DISCUSSION

We have used unnatural amino acid mutagenesis to map key residues in and adjacent to ECL2 of CLR that contribute to the binding of CGRP. The data suggest that I284 and L291 make contacts in the vicinity of Thr6 and Leu12 of bound CGRP. This binding orientation is consistent with the modeled position of calcitonin in the cryo-electron microscopy structure of the calcitonin receptor.9

A number of the residues that we have identified in this study as contributing to the binding of CGRP have been previously identified as being important for CGRP production of cAMP, based on mutagenesis. We predicted that I284 could interact with one or more CGRP residues C-terminal to the 2–7 disulfide-bonded loop of the peptide, conventionally considered to be the mediator of receptor activation.13 However, using an alanine scan, we have previously found that changes to Val8 and Thr9 of CGRP can alter its efficacy.25 Our current model suggests that Val8 is adjacent to H194 in TM2 and Thr9 is in the proximity of the backbone of ECL2. Our study further extends this by demonstrating that Leu16 in CGRP can influence efficacy, potentially by contacting I32 in the ECD of CLR. Furthermore, we have observed during a replica exchange Monte Carlo simulation of the bound CGRP (to investigate dynamic stability), L16 periodically interacts with V198 (ECL1) and I284 of CLR. V198 and I32 have both been shown to be important for cAMP production by CGRP.40,41 These observations suggest how CGRP residues outside of the classic 1–7 “activation” domain of the peptide can influence efficacy; we speculate that Leu16 may be particularly important in interacting with both the ECLs and ECD of CLR. It may therefore be possible to develop truncated analogues that retain agonist activity by interacting with the upper regions of the TM bundle and ECLs, as has been done for GLP-1.42

Within the 1–7 region of CGRP, our previous structure–activity work indicated that Thr6 was essential for agonism. In fact, this residue is absolutely conserved in all members of the calcitonin family of peptides.25,34 In our model, Thr6 of CGRP is close to not only L291 of TM5 but also H295 a turn lower. Previous mutagenesis has shown that H295 is important for the activation of all CLR/RAMP complexes with both CGRP and adrenomedullin;11 its equivalent is also important for agonism by calcitonin in the calcitonin receptor.9 Molecular dynamics simulations indicated that H295, at the top of TM5, controls interhelical contacts among TM3, TM5, TM6, and TM7 of CLR. A rotation of this residue caused helix sliding and tilting that opened the G protein binding pocket.11 Residues deeper in the receptor that are important for receptor activation such as E34811 would seem unlikely to be significant ligand contacts, unless they are connected by bound water molecules.

Although I284 and L291 were the main contacts revealed in this study, it is interesting to note that there were a range of minor contacts in the distal part of ECL2. The pattern here was that every other residue between I284 and L290 appeared as contacts. Interestingly, our previous mutagenesis and modeling have suggested that this forms a short stretch of β-strand, along the rim of the cavity between TM4 and TM5.12 This secondary structure element is consistent with the labeling pattern.

Figure 6. Molecular model of the CLR (brown ribbon)/RAMP1 (blue ribbon; TM region omitted for the sake of clarity) complex with bound CGRP (pink ribbon). (A) Detail of predicted interactions of CGRP with I284 and L291. (B) Predicted position of bound CGRP with residues on the periphery of the binding site. CLR residues shown in space-filling mode; CGRP residues colored green. The models can be viewed online using Molecule Viewer at https://bit.ly/2FnsQAHp (recommended browsers are Google Chrome in incognito mode or Firefox). (C) Full length structure of the CLR complex. The positions of TM5 and helix 8 are highlighted to aid orientation.

Figure 7. Comparison of the ability of [Leu12Ala]- and [Leu16Ala]- human α-CGRP (A) to stimulate cAMP in HEK293S cells transfected with HA-CLR and myc-RAMP1 or (B) to displace [125I]CGRP at the CGRP receptor. Values are means ± SEM; n = 3 or 4.
In eight of 15 of the human family B GPCRs, the equivalent of I284 is glutamic or aspartic acid; in three additional receptors, it is threonine. Only in CLR, the calcitonin receptor (leucine), and the corticotrophin receptors (phenylalanine) is the corresponding residue hydrophobic. Thus, within family B GPCRs, this residue may be important in ligand specificity; it faces into the core of the receptor where it can be resolved in structures. In the calcitonin receptor, the equivalent residue is important for the binding of human and porcine but not salmon calcitonin, indicating how similar ligands can differentially interact with residues within their binding pockets. The absence of labeling for W283 is noteworthy as this residue is highly conserved in family B GPCRs and mutation may show deleterious effects. Its side chain orientation is variable in the current family B GPCR structures, either facing into the core in the calcitonin and glucagon receptor structures or inserting between TM3 and TM4 in other structures. These data suggest that W283 does not contact the peptide, indicating it is most likely to be inserted between TM3 and TM4 (Figure 6B).

The techniques described in this paper may be applied to other receptors. While previously the method has been applied to the GLP-1 receptor, our experience is that significant refinement is required. In the study presented here, it was necessary to illuminate the receptor for 1 h to obtain cross-linking, rather than for the 5 min used in the GLP-1 study. This difference may reflect the fact that the GLP-1 study was focused on the extracellular domain of the receptor; greater exposure times may be necessary for residues within the membrane domain. In studies of CRF1R, where transmembrane residues were also targeted, illumination was performed for 20–40 min. Thus, investigators should optimize labeling conditions to ensure their protocols are maximally effective. The cross-linking efficiency will also depend on the power of the UV source and the distance of the sample from the light, which are likely to vary between laboratories.

In conclusion, this study suggests that CGRP binds to a pocket formed by the ECLs of CLR, penetrating approximately two turns into the TM region. Direct contacts with residues here in the ECLs form the basis of the agonist-induced conformational shift in CLR during receptor activation.

**AUTHOR INFORMATION**

**Corresponding Author**
*E-mail: d.r.poyner@aston.ac.uk. Telephone: +44 (0)121 204 3997.*

**ORCID**
Debbie L. Hay: 0000-0002-9558-5122
Margaret A. Brimble: 0000-0002-7086-4096
David R. Poyner: 0000-0003-1590-112X

**Author Contributions**
The manuscript was written through contributions of all authors. J.S., R.U., J.J.G., and M.L.G. performed experiments. T.P.S. provided constructs for photoaffinity labeling and technical advice. P.W.H. and M.A.B. synthesized peptides. J.S., C.A.R., T.P.S., D.R.P., and D.L.H. interpreted experiments and wrote the paper. All authors have given approval to the final version of the manuscript.

**Funding**
This work was supported by BBSRC grants to D.R.P. (BB/M007529/1) and C.A.R. (BB/M006883/1). C.A.R. is a Royal Society Industry Fellow. D.L.H. is a James Cook Research Fellow of the Royal Society of New Zealand.

**Notes**
The authors declare no competing financial interest. To access the research data supporting this publication, see https://doi.org/10.17036/researchdata.aston.ac.uk.00000368.

**ACKNOWLEDGMENTS**
The authors thank Dr. Casandra Koole and Dr. Harriet Watkins for advice. Furthermore, the authors thank Merry Wang at Autodesk for help and advice with Molecule Viewer.

**ABBREVIATIONS**
aZP, azido l-phenylalanine; AM, adrenomedullin; CGRP, calcitonin gene-related peptide; CLR, calcitonin receptor-like receptor; CT, calcitonin; Dde, 1-(4,4-dimethyl-2,6-dioxacyclohexylidene)ethyl; DMF, dimethylformamide; ECD, extracellular domain; ECL, extracellular loop; GLP-1, glucagon-like peptide-1; GPCR, G protein-coupled receptor; RAMP, receptor activity-modifying protein; TM, transmembrane.

**REFERENCES**
Biochemistry


(20) Cohn, I., Katritich, V., Sun, T., Xiang, Z., Siu, F. Y., Beyermann, M., Stevens, R. C., and Wang, L. (2013) Genetically encoded chemical probes in cells reveal the binding path of urocrtin-1 to CRF class B GPCR. Cell 155, 1258–1269.


