The μ opioid receptor crystal structure with BU72 is a

2 covalent adduct

- 3 Thomas A. Munro
- 4 School of Life and Environmental Sciences, Deakin University, Melbourne, Australia
- 5 Current email: thomas@munro.com
- 6 <u>orcid.org/0000-0002-3366-7149</u>

7 Abstract

8 Background

9 The first crystal structure of the active μ opioid receptor (μOR) exhibited several 10 unexplained features. The ligand BU72 exhibited many extreme deviations from ideal 11 geometry, along with unexplained electron density around the benzylic carbon. I 12 previously showed that inverting the benzylic configuration resolved these problems, 13 establishing revised stereochemistry of BU72 and its analog BU74. However, another 14 problem remains unresolved: additional unexplained electron density contacts both 15 BU72 and a histidine residue in the N-terminus.

16 Results

Here I show that these short contacts and uninterrupted density are inconsistent with non-covalent interactions. Therefore, BU72 and µOR form a covalent adduct through an unmodeled atom, and the published model as two separate entities is incorrect. A subsequently proposed magnesium complex is also inconsistent with multiple lines of evidence. However, oxygen fits the unexplained density well. While the structure I propose is tentative, similar oxygen-bridged adducts have been reported previously in the presence of reactive oxygen species. Moreover, known sources of reactive oxygen species were present: HEPES buffer, nickel ions, and a sequence motif that forms redox-active nickel complexes. This motif contacts the unexplained density. The adduct exhibits severe strain, and the tethered N-terminus forms contacts with adjacent residues. These forces, along with the nanobody used as a G protein substitute, would be expected to influence the receptor conformation. Consistent with this, the intracellular end of the structure differs markedly from subsequent structures of active μ OR bound to G_i protein.

31 Conclusions

Later G_i-bound structures are likely to be more accurate templates for docking and molecular dynamics simulations of active μ OR. The possibility of reactions like this should be considered in the choice of protein truncation sites and purification conditions, and in the interpretation of excess or unexplained density.

36



Keywords

40 BU72; covalent adduct; crystal structure; μ opioid receptor; revised stereochemistry

42 Background

BU72 is a μ opioid of exceptionally high affinity and potency (Figure 1) [1, 2]. Its dissociation constant (*K*_i) for μ OR ranges from 0.15 nM in crude brain membranes [1], to lower values in transfected cell membranes [2, 3], and as low as 0.01 nM for purified μ OR with G_i protein [3]. Very few ligands for any protein exceed this extraordinary affinity, which is considered an effective upper bound on the strength of non-covalent binding [4].



49

50 Figure 1: Structures of BU72 and analogs

51

⁵² BU72 was the ligand in the first crystal structure of active μ OR [3]. As noted there, the ⁵³ electron density exhibited two unexplained features. Firstly, fitting the published ⁵⁴ structure of BU72 (**1a**, Figure 1) required a near-planar orientation of the phenyl group, ⁵⁵ an implausibly high-energy conformation that required many extreme deviations from ⁵⁶ ideal geometry and left unexplained density around the benzylic carbon (Figure 2a). ⁵⁷ The authors considered the possibility that the ligand was actually imine **2** (Figure 1), ⁵⁸ whose planar sp² benzylic carbon would resolve this problem, but this was not detected in mass spectra of the crystallization mixture [3]. In a preprint, I proposed an alternative: a revised structure for BU72 with the phenyl group in the opposite (*R*) configuration (**1b**, Figure 1) [5]. Revised structure **1b** fits in a low-energy conformation, eliminating the geometric outliers and unexplained density around the phenyl group, and yielding superior validation metrics (Figure 2b) [5]. The similar binding affinities of BU72 and imine **2** [1] provide further support to structure **1b**: the equatorial phenyl group in **1b** is approximately planar, as in imine **2**, which also fits the density in a low-energy conformation, unlike **1a** (see Extended Data Fig. 4e in [3]).



Figure 2: Phenyl group geometric outliers and unexplained electron density for original
(1a) and revised (1b) structures of BU72. Colors: fitted structures (*black*); ideal
structures (*grey*); geometric outliers in the phenyl group (Z scores, *red*); 2*F*o-*F*c density
(2.5σ, *violet*); *F*o-*F*c omit density (2σ, *green*). Adapted from [5]

73

The original proposed configuration of **1a** was based on unpublished nuclear Overhauser effect (nOe) data, and the basis for the necessary NMR assignments was not stated [<u>1</u>, <u>2</u>]. Thus, no published data support the original assignment, and the structure of BU72 should be revised to **1b**. The authors of the crystal structure, including the lead author of the original synthesis, accepted this revision in a correction notice [<u>6</u>]. However, the revised structure was not shown. Protein Data Bank entry

5C1M was also corrected (version 2.0). Note that the structure of the analog BU74 (**3**, Figure 1) should also be revised, since they differ only in the N-substituent [7]; their synthetic routes diverge after establishment of the phenyl configuration, and the benzylic hydrogen is not exchangeable.

However, a second puzzling feature of the crystal structure remains unexplained after this revision. The truncated N-terminus of the receptor, which is highly disordered and hence unresolved in other opioid receptor structures, unexpectedly intrudes into the binding pocket [3]. The third residue, His54, clashes with BU72. The overlapping atoms also contact a pocket of strong, unexplained electron density (Figure 3). The atom responsible for this density could not be identified; experiments testing for an alternative ligand structure or a coordinated heavy metal were unsuccessful [3]. The atom was ultimately omitted from the model altogether. The revised model with **1b** (5C1M v.2) reduced the clash between ligand and receptor, but did not account for the unexplained density. All results in this paper are based on revised structure **1b**.



Figure 3: Clashes and unexplained density between BU72 (1a) and His54 in the
original model (5C1M v.1.5). 2Fo-Fc density (*blue*) and Fo-Fc omit density (*green*) are
shown at the indicated levels. Clashing N atoms are shown as spheres

99

Other authors later proposed that the missing atom is a magnesium ion [8]. This fitted the unexplained density well, while lithium, sodium, nickel, and zinc ions did not [8]. Bond lengths were not given, but were reportedly consistent with a magnesium coordination complex [9].

Results and Discussion

105 The missing atom is not magnesium

I first refined a complex with the previous candidate, Mg^{2+} . Consistent with the earlier reports [8, 9], this gave a good fit, with no excess or unexplained density above 2.5 σ (Figure 4; data in Additional files 1 and 2). However, contrary to the prior reports, the

N-Mg bonds were unrealistically short (1.9 and 1.7 Å). Compare the N-Mg bond lengths in structures of subatomic resolution: 2.19 ± 0.06 Å [10]. These bonds are thus 110 extreme outliers, with Z scores of -5 and -9, respectively. The high resolution of the 111 structure (2.1 Å) allows strong conclusions about bond lengths, with a diffraction 112 precision index (DPI) of 0.22 Å for the Mg²⁺ ion [11]. Note also that the ion is not 113 centered in the density even with these unrealistically short distances, suggesting that 114 the actual bonds must be shorter still (Figure 4). This resulted in a poor real-space R 115 value (RSR) of 0.32 for the Mg²⁺ ion, despite good values for His54 (0.11) and BU72 116 117 (0.08).



- 118
- Figure 4: Proposed magnesium complex [8], with bond lengths and B-factors (*red*)
- 120

121 A later report from the same group added a third bond to the model [9], from Mg²⁺ to 122 Tyr148^{3x33} (using GPCRdb numbering [12]) (Figure 5). However, this would require an 123 O–Mg bond length of 3.1 Å; compared with high-resolution structures (2.10 ± 0.04 Å), 124 this is untenable (Z = 25) [10]. It is instead suggestive of a hydrogen bond to another element. Note also the large gap in the electron density along this proposed bond,
unlike the strong and uninterrupted density for the bonds to BU72 and His54 (Figure
5). Additionally, note the highly asymmetrical geometry required, with a bond angle of
105°, compared to 90° for the N atoms; magnesium complexes are symmetrical [10].



129



131

Other evidence against Mg²⁺ was revealed by CheckMyMetal [<u>13</u>]. The values of five of the eight parameters evaluated were classed as dubious, including three that strongly suggest a misidentified element:

• A much higher temperature factor (B-factor) for the ion than the bonding partners (Figure 4); since bonds transmit thermal motion, this is implausible [14].

- Bonding to an amine, which is positively charged at this pH (7.5), while Mg²⁺
 favors neutral or negatively-charged bonding partners [15].
- An incomplete coordination sphere. The expected number of bonds is six, or in
 rare cases four or five; a value of two is extremely rare in high-resolution
 structures [16].

While it could be speculated that unresolved water molecules complete the coordination sphere, this is implausible since the rest of the complex is resolved with full occupancy, as are many structured water molecules elsewhere in the binding pocket [3].

Finally, no source of magnesium is mentioned in the experimental method [3]. Collectively, the above lines of evidence firmly exclude Mg²⁺ as a candidate.

The missing atom forms covalent bonds to both BU72 and His54

While the element is evidently misidentified, the fit of the Mg²⁺ ion to the density does firmly establish a non-hydrogen atom in this approximate position. As noted above, this missing atom is likely nearer to both His54 and BU72 than the modelled position of Mg²⁺; that is, < 1.9 Å from each (Figure 4). This is much too close for non-covalent interactions (\geq 2.4 Å) [17], which would also not result in strong, uninterrupted electron density connecting the three atoms. For instance, the protonated tertiary amine of BU72 forms a charge-assisted hydrogen bond (salt bridge) to aspartate Asp147^{3x32} (Figure 6); these are among the shortest of all noncovalent interactions [17]. Nonetheless, the N···O distance is 2.6 Å, and the regions of high electron density are widely separated, in striking contrast to the continuous density surrounding the purported Mg²⁺ complex (Figure 6). Therefore, the unidentified atom is covalently bonded to both BU72 and µOR; that is, they form an adduct.

161



162

Figure 6: Electron density comparison of the proposed Mg²⁺ complex with the salt bridge to Asp147^{3x32}

165

While this evidence does not establish the identity of the missing atom, it does establish that the published model of BU72 and the receptor as discrete entities is incorrect. A model of the adduct with the bridging atom left unidentified would be correct, albeit incomplete; hundreds of Protein Data Bank (PDB) structures contain unidentified atoms (ligand UNX). Below I consider other candidates.

The missing atom is very unlikely to be a metal, but may be oxygen

The CheckMyMetal validation report for the magnesium complex suggested alternative metals as better candidates: copper, iron, cobalt, nickel, manganese, and zinc. However, each of these also gave multiple outliers when validated. Also, of these metals, only nickel was present during preparation of the crystals (in the affinity column used for purification) [3]. The bond lengths are more plausible than for magnesium, since N–Ni bonds are short (1.88 ± 0.03 Å) [10]. However, as noted above, nickel did not fit the electron density, leaving a substantial excess [8]; further evidence against nickel and other heavy metals is the lack of anomalous scattering noted in the originalreport [3].

The only metal in the buffer solution, sodium, also gave five CheckMyMetal outliers, including even more extreme outliers from typical N–Na bond lengths (2.46 ± 0.02 Å, Z = -29 and -40) [10], and a much worse fit to the density than magnesium [8]. Indeed, no metal forms coordination bonds to N shorter than 1.76 Å [10]. It is thus extremely implausible that the missing atom is a metal.

Given the above, it appears that the missing atom is a non-metal approximately isoelectronic with magnesium, but that forms shorter bonds. The element must also be at least divalent, and can probably form hydrogen bonds given its distance to Tyr148^{3x33} (~3.1 Å). One candidate meeting these criteria is oxygen; based on electron density alone, water molecules are frequently misidentified as magnesium [18].

A known source of reactive oxygen species contacts the unexplained

192 density

Formation of an oxygen-bridged adduct between the secondary amine of BU72 and the imidazole ring of His54 would require harsh conditions. Reactive oxygen species (ROS), for instance, can oxidize secondary amines [19] and histidine [20]. But how might these arise? Surprisingly, several potential sources of ROS were present. The BU72-μOR complex was purified and crystallized in HEPES buffer, which generates hydrogen peroxide on exposure to light [21]. HEPES has also been reported to enhance metal-catalyzed generation of other ROS from hydrogen peroxide [22]. A further potential source is the N-terminus, which contains a sequence motif known to generate ROS. The N-terminus used was truncated, leaving glycine as the first residue and histidine as the third [3]. This sequence motif (Gly-Xaa-His) forms redox-active nickel coordination complexes [23]. Moreover, a nickel affinity column was used for

purification [3], and the Gly-Xaa-His motif can capture Ni²⁺ ions from these columns
[24-26]. The resulting square planar nickel complexes catalyze the decomposition of
hydrogen peroxide to other ROS such as hydroxyl radicals [23]. Thus, the conditions
used were sufficient to generate ROS immediately adjacent to His54, potentially
oxidizing both the residue itself and BU72.

A search of PDBeMotif [27] revealed eight protein structures in which square planar Gly-Xaa-His-Ni²⁺ complexes were resolved: PDB entries 1JVN, 1XMK, 2RJ2, 3RDH, 3UM9, 3ZUC, 4I71, and 4OMO. In three cases, the nickel was not added during crystallization, but unexpectedly captured during affinity chromatography: 1JVN [24], 3UM9 [25], and 3ZUC [26]. Intriguingly, in 1JVN the electron density was not consistent with the expected ligand structure; no density supported several of the atoms, suggesting partial decomposition [24]. The buffer used, PIPES, is an analog of HEPES that also generates hydrogen peroxide [28] and other ROS [22]. This provides a plausible explanation for the decomposition of the ligand.

Proposed structure of an oxygen-bridged adduct

Two previous reports of adduct formation between aminoxyl radicals and imidazole rings are shown in Figure 7a [20, 29]. These suggested potential structure **6** for an adduct between BU72 and His54 (Figure 7b). The stereochemistry of the bond to the modified histidine residue was dictated by the observed density. A possible intermediate aminoxyl radical is also shown; these can form from oxidation of secondary amines by ROS [19].



b) adduct proposed here, with possible intermediate



225

Figure 7: **a)** Reported adducts **4** ([29], Scheme 2) and **5** ([20], Figure 7c). **b)** Adduct **6** proposed here, with the nickel complex and a possible aminoxyl intermediate

228 Oxygen-bridged adduct 6 fits the unexplained density

Substituting adduct **6** for His54 and BU72 gave an excellent fit, with no excess or unexplained density even at 2σ (Figure 8; data in Additional files 3-6). Both bonds to oxygen were of typical length (1.5 Å), and were resolved up to 4.2σ – that is, higher density than most of the ligand itself and surrounding side-chains. Unlike Mg²⁺, the oxygen atom was well centered in the density. Oxygen also gave a superior B-factor to Mg²⁺, both lower and consistent with its bonding partners, making this a much more plausible candidate element (Figure 8) [14]. The lower B-factor for oxygen results in a more precise fit (DPI 0.14 vs 0.22 Å). Indeed, it is among the most precisely-resolved atoms in the entire structure, which is itself the highest-resolution structure of μ OR to date. The bridging oxygen and modified histidine moiety make favorable polar contacts with Tyr148^{3x33}, which are close to the length of a weak hydrogen bond.



240

Figure 8: Fit of adduct **6** to density, with B-factors (*red*) and polar contact distances to

- 242 Tyr148^{3x33}
- 243

The adduct is highly strained

The bound geometry of adduct **6** gave acceptable validation metrics, which were superior to the original model of BU72, **1a** (Table 1; data in Additional file 7).

247

- Table 1: Ligand validation: geometry relative to GRADE restraints, and fit to electron
- 250 density from PDB validation reports

PDB Structure	5C1M (v1.5)	8E0G				
Ligand	BU72 (1a)	adduct 6				
Geometry						
Geometric outliers $(Z > 2)$	26	10				
Severe outliers $(Z > 5)$	9	1				
Bond angle root mean square Z (RMSZ)	3.23	1.52				
Bond length root mean square Z (RMSZ)	3.32	1.13				
Fit to electron density						
Real-space correlation coefficient (RSCC) ^a	0.914	0.951				
Real-space R (RSR)	0.090	0.081				
al ower values are better except for RSCC						

^a Lower values are better except for RSCC

252

The only severe outlier was the bond angle at the bridging oxygen (131° vs the ideal, 109°: Z = 7.2). There are several indications that this is real strain rather than a fitting artefact, however. The angle is clearly resolved at high density, and is consistent with tension from the tethered N-terminus. The phenyl group is bent 11° out of plane, consistent with being pulled against the adjacent residue Ile144^{3x29} by the same tension (Figure 9). This bend is also clearly resolved, and is comparable to those seen in severely strained aromatic residues at subatomic resolution [30]. It also yields a more complementary fit to Ile144^{3x29} than the original model, as well as eliminating another small pocket of unexplained density (Figure 9).



Figure 9: Fit of phenyl group to adjacent residue Ile144^{3x29}, shown with solventaccessible surfaces (**a**: original model (5C1M v.1.5); **b**: adduct)

- 265
- Strain is also evident in the N-terminus itself: in both this model and the original (5C1M
- v.1.5), Thr60 adopts a rare and high-energy *cis*-peptide bond, and there are many
- 268 energetically unfavorable clashes along the peptide backbone (Figure 10).



Figure 10: Polar contacts (< 3.6 Å) and clashes of the tethered Nterminus in the adduct
 model. Note the high-energy *cis*-peptide bond at Thr60

Alternate modelling can eliminate the *cis*-peptide bond, as in the revised version of the original model (5C1M v.2). However, this results in a worse fit to the density, which is extremely weak in this region: several side-chains and even parts of the backbone are unresolved at 1 σ , yielding eight RSR outliers in the N-terminus, five of which are severe (Figure 11). Atomic displacements in the N-terminus are also extremely high: the occupancy-weighted average B-factor (OWAB) of the last seven residues (58-64) are higher than 95% of residues in the structure. Indeed, Gln59 has the highest value in the entire structure, 159 Å², compared to a median of 46. The above features (poor density coverage, high B-factors, clashes and a probable *cis*-peptide bond) imply that the N-terminus is constrained in an extremely unfavorable high-energy state by the tethered ligand.



Figure 11: The N-terminus in the revised version of the original model (5C1M v.2), colored by B-factor. Note poor electron density coverage for some residues; RSRZ scores > 5 (severe outliers) are given in brackets

289

Despite the very strong interactions apparent between BU72 and His54, removal of the side chain of His54 by receptor mutagenesis had no detectable effect on the affinity or potency of BU72 [3]. This seeming paradox, however, is consistent with the mechanism proposed here. Since the full-length receptor was used for the assays, the Gly-Xaa-His motif was not at the N-terminus, and therefore nickel complexation and adduct formation could not occur. Thus, binding would be unaffected by the presence or absence of His54.

Adduct strain, N-terminal contacts, and nanobody Nb39 distort the receptor, confounding inferences about the active conformation

The forces required to tether the ligand and N-terminus in high-energy conformations must affect the rest of the receptor. Compounding this, the N-terminus makes numerous strong contacts throughout the binding pocket, including a dense network of

polar contacts and clashes with transmembrane helices and extracellular loops (Figure10).

In addition to the strain in the N-terminus and the contacts it makes, another factor likely to influence the receptor conformation is the intracellular binding partner used, the G protein mimetic nanobody Nb39. Nanobodies are known to yield slightly different receptor conformations than naturally-occurring G proteins [31].

The largest movement during activation of G protein-coupled receptors (GPCRs) involves TM6. Viewed from the intracellular end, TM6 pivots outwards and rotates clockwise; this 'macroswitch' occurs in all GPCRs studied to date [31, 32]. This shift is markedly different in the BU72- μ OR-Nb39 structure than in later structures of active μ OR bound to G_i protein (Figure 12 and Table 2). Although these 13 later structures feature diverse μ opioids bound to mouse or human μ OR, they cluster very tightly in this key region. The BU72-bound structure is a clear outlier, with TM6 much closer to TM5, and rotated in the opposite direction. As a result, intracellular loop 3 (ICL3) bunches outwards in a disordered loop, rather than being pulled into a helix as in the G_i-bound structures. These differences appear to be largely due to Nb39, since the structure of the κ opioid receptor (κ OR) bound to the same nanobody is similar (Figure 12).



- Figure 12: Overlay of TM5, TM6, and ICL3 (inactive, Nb39-bound, and Gi-bound). See
- 322 Table 2 for PDB identifiers and other details

323

324 Table 2: Opioid receptor structures discussed

PDB	B OR		OR Species		Bound to		to	Ligand	Ref.
	μ	K	mouse	human	Nb39	Gi	_		
5C1M	•		•		•			BU72	[<u>3</u>]
6B73		•	•		•			MP1104	[<u>33</u>]
6DDF	•		•			•		DAMGO	[34]
7SBF	•		•			•		PZM21	[<u>35</u>]
7SCG	•		•			•		FH210	[<u>35</u>]
7T2G	•		•			•		mitragynine pseudoindoxyl	[<u>36</u>]
7T2H	•		•			•		lofentanil	[<u>36</u>]
7U2K	•		•			•		C6 guano	[<u>37</u>]
7U2L	•		•			•		C5 guano	[<u>37</u>]
7UL4	•		•				•	alvimopan (inverse agonist)	[<u>38</u>]
8EF5	•			•		•		fentanyl	[<u>39</u>]
8EF6	•			•		•		morphine	[39]
8EFB	•			•		•		oliceridine	[<u>39</u>]
8EFL	•			•		•		SR-17018	[<u>39</u>]
8EFO	•			•		•		PZM21	[39]
8EFQ	•			•		•		DAMGO	[<u>39</u>]

Another conspicuous discrepancy between the BU72-bound structure and the others is in helix 8 (H8). Activation of class A GPCRs, such as opioid receptors, involves an inward shift of H8, making and breaking contacts at its base (see Figures 3 and 4 in [32]). Relative to the inactive structure, the base of H8 shifts noticeably more in BU72- μ OR-Nb39 than in the μ OR-G_i structures or κ OR-Nb39, which again cluster tightly (Figure 13). This suggests that the nanobody itself is not responsible for the discrepancy, but rather some other factor, such as distortion due to strain in the adduct.



333

- Figure 13: Overlay of H8 (inactive, Nb39-bound, and Gi-bound). See Table 2 for PDB
 identifiers and other details
- 336

Whether due to the influence of the adduct, the nanobody or both, these differences from the µOR-G_i structures are experimental artefacts, and the consistency between the G_i-bound structures establishes them as superior templates for modeling the active conformation.

341 Proposed experimental tests of adduct formation

In the original study, a search for alternative ligands to account for the unexplained density was unsuccessful. The mass spectrum of the crystallization mixture revealed a molecular ion consistent with BU72, but no others of similar mass [3]. However, the intact adduct would not be detectable in solution, and one decomposition product per binding site would yield negligible concentrations relative to saturating BU72. An alternative test would be for modification of His54: proteolysis of the receptor and mass spectrometry of the fragments should reveal either the adduct or decomposition products. A simpler alternative would be to substitute a short Gly-Xaa-His-containing peptide for the receptor, although this might also result in side-reactions. The initial nickel complex itself should be detectable spectroscopically, and may indeed give a noticeable yellow color to the solution [23].

An obstacle to isolation of the adduct may be instability. Previously-reported adducts **4** and **5** were not isolated, but detected only by mass spectrometry as reaction intermediates [20, 29]. However, the tethered conformation of the N-terminus separates Gly52 from His54, rendering a nickel complex between the two residues impossible (Figure 11). Thus, adduct formation would liberate the ion and end the catalytic cycle. Moreover, the 'lid' formed by the N-terminus almost entirely occludes the binding pocket [3], leaving only a narrow tunnel filled with structured water molecules. Thus, the adduct bonds are sterically shielded, which may inhibit further reactions.

362 Wider implications, and precautions against ROS generation

The risk of unexpected complexes and oxidations like this is not specific to the structures discussed here. The conditions that led to these reactions, in both this case and previously [24], are widely used. Many common methods for the cleavage of fusion

proteins (thrombin, factor Xa, tobacco etch virus protease, and rhinovirus 3C protease) leave glycine as the N-terminal residue [40]. Unsurprisingly then, the N-terminal Gly-Xaa-His motif is common in the Protein Data Bank, appearing in >7,000 sequences (~4% of the total). Nickel affinity columns are also widely used. Many of these proteins would therefore be expected to form Gly-Xaa-His-Ni²⁺ complexes. However, the first few residues of the N-terminus are almost invariably disordered: 97% of human proteins have disordered terminal residues [41], and 42% of all disordered residues are in the N-terminus [42]. Thus, these complexes are very unlikely to be resolved, and are therefore likely to go undetected. Peroxide-generating buffers such as HEPES are also ubiquitous; thus, quite common procedures for protein preparation inadvertently generate ROS. Oxidation by ROS can have many undesirable effects on proteins, from modifying side chains (which may influence the overall conformation) to cleaving the amide backbone [43].

The possibility of reactions like this should be considered in the choice of truncation sites and purification conditions for protein isolation. Generation of nickel complexes, ROS, and subsequent reactions could be prevented by choosing a different cleavage site (with a third residue other than histidine) or a nickel-free purification method. Where a nickel complex is desired, for instance to promote crystallization [24] or assist in phasing [26], a non-piperazine buffer such as Tris or MES could be used to avoid or reduce ROS generation [44].

Conclusions

In summary, the density observed between BU72 and His54 is not consistent with noncovalent interactions or a metal coordination complex, and must instead represent covalent bonds to a non-metal atom, approximately isoelectronic with Mg²⁺. The density firmly establishes the presence of this atom and two covalent bonds, and suggests a polar contact with Tyr148. While this evidence does not unambiguously identify the atom, it does establish that the published model is incorrect. The use of conditions known to generate ROS, along with adducts reportedly previously in the presence of ROS, suggest a tentative structure and mechanism for the formation of an oxygen-bridged adduct. All features examined are consistent with this proposal.

The structure differs in several respects from subsequent structures of μ OR bound to G_i protein, likely due to the use of a nanobody, severe strain within the N-terminus, and its contacts with surrounding residues. These subsequent μ OR-G_i structures are likely to be more accurate templates of the active receptor for docking and simulations of molecular dynamics. Oxidative artefacts like this can be prevented by careful choice of truncation sites and purification conditions.

402 Methods

Starting from the previously reported model [5] of μOR with 1b, Mg²⁺ was added to the
center of the unexplained density with sphere refinement using Coot [45] in CCP4i2
[46], and uploaded with the original structure factors to PDB-REDO server [47] for
automated refinement. The resulting complex was submitted to CheckMyMetal [13] for
validation; all suggested alternative metals were also resubmitted for validation.

The ideal structure and geometric restraints of the 1b-histidine adduct 6 were 408 generated using GRADE server [48]. BU72 was deleted from the original model, His54 409 was mutated to the adduct, and the model fitted and refined as above. Because the 410 PDB validation report did not evaluate the geometry of adduct 6, ligand distortions in 411 the bound ligands were tabulated in Coot and used to calculate Z scores, comparing 412 ideal values and standard deviations from GRADE with modeled values for 1a, 1b and 413 6 (Additional file 7). Diffraction precision indexes were calculated using Online DPI 414 [11]. Protein structures were aligned and illustrated using Pymol [49], and annotated 415

using Inkscape [50]. Small-molecule structural formulae were drawn using
Marvinsketch [51], and are provided in Chemical Markup Language as Additional file
8.

419 **Declarations**

- 420 Ethics approval and consent to participate
- 421 Not applicable.
- 422 Consent for publication
- 423 Not applicable.

424 Availability of data and materials

- 425 Coordinates and structure factors for the adduct have been deposited in the Protein 426 Data Bank (under accession number <u>8E0G</u>). These and other datasets supporting the 427 conclusions of this article are included in the supplementary information files. An 428 interactive comparison of the adduct and original model, including electron density, is
- 429 available at: molstack.bioreproducibility.org/p/Y7FU
- 430 **Competing interests**
- 431 I declare no competing interests.
- 432 Funding
- 433 Not applicable.
- 434 Authors' contributions
- 435 I am the sole author.

436 Acknowledgements

Robbie Joosten kindly modified the code of PDB-REDO server to enable refinement ofthe adduct.

439 Supplementary Information

440	Additional file 1: BU72-Mg-muOR model.cif
441	Coordinates of the BU72-Mg ²⁺ - μ OR complex
442	
443	Additional file 2: BU72-Mg-muOR phases.mtz
444	Structure factors of the BU72-Mg ²⁺ - μ OR complex
445	
446	Additional file 3: BU72-muOR adduct model.cif
447	Coordinates of the BU72- μ OR adduct
448	
449	Additional file 4: BU72-muOR adduct phases.mtz
450	Structure factors of the BU72- μ OR adduct
451	
452	Additional file 5: BU72-histidine adduct ideal structure.pdb
453	Ideal coordinates of BU72-histidine adduct 6 (GRADE server)
454	
455	Additional file 6: BU72-histidine adduct restraints.cif
456	Geometric restraints of BU72-histidine adduct 6 (GRADE server)
457	
458	Additional file 7: GRADE ligand outliers.xlsx
459	Geometric outliers of BU72 and the BU72- μ OR adduct (GRADE server)
460	
461	Additional file 8: chemical structures.cml
462	Chemical structures (structural formulae) of the small molecules

⁴⁶³ **References**

464

Husbands SM, Lewis JW. Morphinan cyclic imines and pyrrolidines containing
 a constrained phenyl group: high affinity opioid agonists. Bioorg Med Chem Lett.
 1995;5(24):2969-74. doi.org/cbmj9g

468 2. Neilan CL, Husbands SM, Breeden S, Ko M-C, Aceto MD, Lewis JW, et al. 469 Characterization of the complex morphinan derivative BU72 as a high efficacy, long-470 lasting μ -opioid receptor agonist. Eur J Pharmacol. 2004;499(1):107-16. doi.org/fpdvz9

471 3. Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn 472 AL, et al. Structural insights into μ -opioid receptor activation. Nature. 473 2015;524(7565):315–21. doi.org/f7m78z

474 4. Smith RD, Engdahl AL, Dunbar JB, Carlson HA. Biophysical limits of protein-475 ligand binding. J Chem Inf Model. 2012;52(8):2098-106. <u>doi.org/f36nns</u>

476 5. Munro TA. Revised (β-phenyl) stereochemistry of ultrapotent μ opioid BU72.
477 bioRxiv. 2020. doi.org/dq7s

Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn
AL, et al. Author correction: Structural insights into μ-opioid receptor activation. Nature.
2020;584(7820):E16. doi.org/d5jj

Husbands SM, Neilan CL, Broadbear J, Grundt P, Breeden S, Aceto MD, et al.
BU74, a complex oripavine derivative with potent kappa opioid receptor agonism and delayed opioid antagonism. Eur J Pharmacol. 2005;509(2):117-25. doi.org/drkbg6

8. Chan HCS, Xu Y, Tan L, Vogel H, Cheng J, Wu D, et al. Enhancing the signaling of GPCRs via orthosteric ions. ACS Cent Sci. 2020;6(2):274-82. <u>doi.org/dk25</u>

486 9. Zou R, Wang X, Li S, Chan HCS, Vogel H, Yuan S. The role of metal ions in G
487 protein-coupled receptor signalling and drug discovery. WIREs Comput Mol Sci.
488 2021;12(2):e1565. doi.org/gmcwd9

Kuppuraj G, Dudev M, Lim C. Factors governing metal–ligand distances and
 coordination geometries of metal complexes. J Phys Chem B. 2009;113(9):2952-60.
 <u>doi.org/dqgnpi</u>

492 11. Kumar KSD, Gurusaran M, Satheesh SN, Radha P, Pavithra S, Thulaa
493 Tharshan KPS, et al. Online_DPI: a web server to calculate the diffraction precision
494 index for a protein structure. J Appl Crystallogr. 2015;48(3):939-42. doi.org/f7fbq6

12. Isberg V, de Graaf C, Bortolato A, Cherezov V, Katritch V, Marshall FH, et al.
Generic GPCR residue numbers – aligning topology maps while minding the gaps.
Trends Pharmacol Sci. 2015;36(1):22-31. <u>doi.org/f62gpv</u>

Gucwa M, Lenkiewicz J, Zheng H, Cymborowski M, Cooper DR, Murzyn K, et
 al. CMM – an enhanced platform for interactive validation of metal binding sites. Protein
 Sci. 2022:e4525. doi.org/jpxj

14. Masmaliyeva RC, Babai KH, Murshudov GN. Local and global analysis of macromolecular atomic displacement parameters. Acta Crystallogr Sect D: Biol Crystallogr. 2020;76(10):926-37. <u>doi.org/jght</u>

15. Zheng H, Chordia MD, Cooper DR, Chruszcz M, Muller P, Sheldrick GM, et al.
 Validation of metal-binding sites in macromolecular structures with the CheckMyMetal
 web server. Nat Protoc. 2014;9(1):156-70. doi.org/dpcs

16. Dudev M, Wang J, Dudev T, Lim C. Factors governing the metal coordination
 number in metal complexes from Cambridge Structural Database analyses. J Phys
 Chem B. 2006;110(4):1889-95. <u>doi.org/dbw84h</u>

- Kruse H, Sponer J, Auffinger P. Comment on "Evaluating unexpectedly short
 non-covalent distances in X-ray crystal structures of proteins with electronic structure
 analysis". J Chem Inf Model. 2019;59(9):3605-8. doi.org/d8jz
- 18. Leonarski F, D'Ascenzo L, Auffinger P. Mg²⁺ ions: do they bind to nucleobase
 nitrogens? Nucleic Acids Res. 2016;45(2):987-1004. <u>doi.org/f9sg7d</u>
- Aurich HG. Nitroxides. In: Patai S, Rappoport Z, editors. Nitrones, Nitronates
 and Nitroxides. Patai's Chemistry of Functional Groups. New York, NY: John Wiley &
 Sons; 1989. p. 313-70. doi.org/cp3635
- 20. Ihara H, Kakihana Y, Yamakage A, Kai K, Shibata T, Nishida M, et al. 2-Oxohistidine–containing dipeptides are functional oxidation products. J Biol Chem.
 2019;294(4):1279-89. doi.org/d7n6
- Masson J-F, Gauda E, Mizaikoff B, Kranz C. The interference of HEPES buffer
 during amperometric detection of ATP in clinical applications. Anal Bioanal Chem.
 2008;390(8):2067-71. doi.org/fhkf72
- 524 22. Simpson JA, Cheeseman KH, Smith SE, Dean RT. Free-radical generation by 525 copper ions and hydrogen peroxide. Stimulation by Hepes buffer. Biochem J. 526 1988;254(2):519-23. <u>doi.org/jf6h</u>
- 23. Ueda J-i, Ozawa T, Miyazaki M, Fujiwara Y. SOD-like activity of complexes of nickel(II) ion with some biologically important peptides and their novel reactions with hydrogen peroxide. Inorg Chim Acta. 1993;214(1):29-32. <u>doi.org/d8dbw9</u>
- 530 24. Chaudhuri BN, Lange SC, Myers RS, Chittur SV, Davisson VJ, Smith JL. Crystal 531 structure of imidazole glycerol phosphate synthase: a tunnel through a (β/α)₈ barrel 532 joins two active sites. Structure. 2001;9(10):987-97. <u>doi.org/cjmk4z</u>
- 533 25. Chan PWY, Chakrabarti N, Ing C, Halgas O, To TKW, Wälti M, et al. 534 Defluorination capability of L-2-haloacid dehalogenases in the HAD-like hydrolase 535 superfamily correlates with active site compactness. ChemBioChem. 536 2022;23(1):e202100414. doi.org/jdgn
- Yaniv O, Halfon Y, Shimon LJW, Bayer EA, Lamed R, Frolow F. Structure of
 CBM3b of the major cellulosomal scaffoldin subunit ScaA from *Acetivibrio cellulolyticus*. Acta Crystallogr Sect F: Struct Biol Commun. 2012;68(1):8-13.
 <u>doi.org/fx9pds</u>

- 541 27. Golovin A, Henrick K. MSDmotif: exploring protein sites and motifs. BMC 542 Bioinformatics. 2008;9(1):312. <u>doi.org/cs84rz</u>
- 543 28. Kirsch M, Lomonosova EE, Korth H-G, Sustmann R, de Groot H. Hydrogen
 544 peroxide formation by reaction of peroxynitrite with HEPES and related tertiary amines:
 545 implications for a general mechanism. J Biol Chem. 1998;273(21):12716-24.
 546 doi.org/dhdr4m
- 29. Chen J-H, Ahmed W, Li M-H, Li Z-D, Cui Z-N, Tang R-Y. TEMPO-mediated synthesis of *N*-(fluoroalkyl)imidazolones via reaction of imidazoles with iodofluoroacetate. Adv Synth Catal. 2020;362(1):269-76. <u>doi.org/d7v3</u>
- 50 30. Laulumaa S, Kursula P. Sub-atomic resolution crystal structures reveal 51 conserved geometric outliers at functional sites. Molecules. 2019;24(17):3044. 52 doi.org/fn8x
- 553 31. Zhou Q, Yang D, Wu M, Guo Y, Guo W, Zhong L, et al. Common activation 554 mechanism of class A GPCRs. eLife. 2019;8:e50279. <u>doi.org/fpcw</u>
- 555 32. Hauser AS, Kooistra AJ, Munk C, Heydenreich FM, Veprintsev DB, Bouvier M, 556 et al. GPCR activation mechanisms across classes and macro/microscales. Nat Struct 557 Mol Biol. 2021;28(11):879-88. <u>doi.org/imp2</u>
- 558 33. Che T, Majumdar S, Zaidi SA, Ondachi P, McCorvy JD, Wang S, et al. Structure 559 of the nanobody-stabilized active state of the kappa opioid receptor. Cell. 560 2018;172(1):55-67. <u>doi.org/chz8</u>
- 561 34. Koehl A, Hu H, Maeda S, Zhang Y, Qu Q, Paggi JM, et al. Structure of the μ opioid receptor—G_i protein complex. Nature. 2018;558(7711):547—52. <u>doi.org/cqzg</u>
- 35. Gmeiner P, Wang H, Hetzer F, Huang W, Qu Q, Meyerowitz J, et al. Structure based evolution of G protein-biased μ-opioid receptor agonists. Angew Chem Int Ed.
 2022;61(26):e202200269. doi.org/hp35
- 566 36. Qu Q, Huang W, Aydin D, Paggi JM, Seven AB, Wang H, et al. Insights into 567 distinct signaling profiles of the μ OR activated by diverse agonists. Nature Chem Biol. 568 2022. doi.org/jm7j
- 569 37. Faouzi A, Wang H, Zaidi SA, DiBerto JF, Che T, Qu Q, et al. Structure-based 570 design of bitopic ligands for the μ -opioid receptor. Nature. 2022. <u>doi.org/grcq2m</u>
- 38. Robertson MJ, Papasergi-Scott MM, He F, Seven AB, Meyerowitz JG, Panova
 O, et al. Structure determination of inactive-state GPCRs with a universal nanobody.
 Nat Struct Mol Biol. 2022. <u>doi.org/jnhk</u>
- S74 39. Zhuang Y, Wang Y, He B, He X, Zhou XE, Guo S, et al. Molecular recognition
 of morphine and fentanyl by the human μ-opioid receptor. Cell. 2022;185(23):4361-75.
 <u>doi.org/jmjd</u>
- 40. Waugh DS. An overview of enzymatic reagents for the removal of affinity tags. Protein Expr Purif. 2011;80(2):283-93. <u>doi.org/dn7skj</u>

- 41. Pentony MM, Jones DT. Modularity of intrinsic disorder in the human proteome.
 Proteins. 2010;78(1):212-21. <u>doi.org/b5jhhp</u>
- 42. Lobanov MY, Furletova EI, Bogatyreva NS, Roytberg MA, Galzitskaya OV. Library of disordered patterns in 3D protein structures. PLOS Comput Biol. 2010;6(10):e1000958. <u>doi.org/fmk4dn</u>
- 43. Hawkins CL, Davies MJ. Detection, identification, and quantification of oxidative protein modifications. J Biol Chem. 2019;294(51):19683-708. <u>doi.org/gpz2xk</u>
- 44. Grady JK, Chasteen ND, Harris DC. Radicals from "Good's" buffers. Anal Biochem. 1988;173(1):111-5. <u>doi.org/dv6hcg</u>
- 45. Emsley P. Tools for ligand validation in Coot. Acta Crystallogr Sect D: Biol Crystallogr. 2017;73(3):203-10. <u>doi.org/f9txh4</u>
- 46. Nicholls R. Ligand fitting with CCP4. Acta Crystallogr Sect D: Biol Crystallogr. 2017;73(2):158-70. <u>doi.org/f9ps4d</u>
- 47. Joosten RP, Long F, Murshudov GN, Perrakis A. The PDB_REDO server for macromolecular structure model optimization. IUCrJ. 2014;1(4):213-20. <u>doi.org/dmft</u>
- 48. Smart OS, Womack TO, Sharff A, Flensburg C, Keller P, Paciorek W, et al. GRADE version 1.2.19. 2011. <u>grade.globalphasing.org</u>
- 49. DeLano WL, Schrödinger LLC. PyMOL molecular graphics system (version 2.5.0). 2021. <u>pymol.org</u>
- 598 50. Inkscape Developers. Inkscape (version 1.0). 2021. inkscape.org
- 599 51. ChemAxon. MarvinSketch (version 22.11). 2022. <u>chemaxon.com/marvin</u>
- 600