CapChaR: Estimation of Receptor Channel Capacity

Chaitanya Tandon
2022UCS0089
Mohammed Sarim Shamim
2022UCS0096
Aryan Katiyar
2022UCH0043

Supervised by:
Dr. Mithu Baidya
Dr. Sarada Prasad Gochhayat

Abstract

Cell survival requires cell-to-cell and cell-environment communication. Cells possess specialised receptors that sense external cues and communicate through chemical signals. Although biologists use chemical and biological readouts to measure cell-to-cell communication, the magnitude of information transfer is not fully understood. The information theoretic approach could possibly revolutionise the area of cell-to-cell communication and be leveraged for high-end precision diagnostics, for example, distinguishing between cancerous and non-cancerous cells, where there are differences in the magnitude of intercellular communication. This study investigates the information capacity of G protein-coupled receptor (GPCR) signalling pathways using cyclic adenosine monophosphate (cAMP) response data obtained from HEK293 cells stimulated with arginine vasopressin (AVP). The cAMP response of approximately 65,000 HEK293 cells was analysed across 12 distinct ligand concentrations. Statistical descriptors were calculated for logarithmically transformed cAMP responses, and mutual information analysis explored the relationship between input ligand concentration and cAMP response. Channel capacity, representing the upper limit of mutual information, was examined, providing insights into information transmission capabilities. Further investigations at the single-cell level were conducted, and computational analyses were facilitated using a Python script. This methodological framework offers insights into signalling dynamics in cellular responsiveness and highlights the potential of GPCRs as therapeutic targets, particularly in disease states such as cancer.

Keywords: G protein-coupled receptors (GPCRs), cyclic adenosine monophosphate (cAMP), information theory, channel capacity, cellular signalling, HEK293 cells, arginine vasopressin (AVP), mutual information, single-cell analysis, cancer, therapeutic targets

1. Introduction

1.1 Background information about the topic

1.1.1 GPCRs: Versatile Signal Transducers

G protein-coupled receptors (GPCRs) constitute the largest receptor family within the animal kingdom, with approximately 823 GPCRs encoded by the human and other mammalian genomes [1]. Their structural adaptability has enabled GPCRs to function as efficient transducers of diverse signals, including light, small molecules, and lipo glycoproteins, across the plasma membrane. Notably, over half of all approved pharmaceutical drugs target GPCRs or their associated signalling pathways, underscoring the critical importance of comprehending GPCR signalling mechanisms in both biological and medical contexts [2].

Despite decades of research into GPCR signalling, fundamental aspects of signal transduction by these receptors remain incompletely understood. Key areas requiring further investigation include the robustness, redundancy, specificity, signal amplification, and noise suppression mechanisms inherent to GPCR signalling pathways. To address these complexities, information theory, initially proposed by Claude E. Shannon in 1948 to delineate fundamental limits on signal processing and communication operations, offers a valuable framework [3][4][5].

1.1.2 Channel Capacity: Maximising Information Transmission

A central tenet of information theory, channel capacity, characterises the maximal amount of information reliably transmissible by an information-transmitting system within a given timeframe. Channel capacity is quantified in bits, with a capacity of one bit indicating the system's ability to reliably transmit a binary signal (on or off), while a capacity of n bits resolves 2^n distinct values reliably.

Quantification of mutual information and channel capacity is indispensable for elucidating the stimulus discrimination process inherent to intracellular signalling pathways. However, accurate estimation of probabilities from limited experimental measurements necessitates specialised methodologies, which will be expounded upon in subsequent sections.

1.1.3 Vasopressin Type 2 Receptor (V2R): An Essential GPCR

The vasopressin type 2 receptor (V2R), a prototypical GPCR, activates adenylate cyclase upon ligand binding, leading to an increase in intracellular cyclic adenosine monophosphate (cAMP) levels. V2R plays a pivotal role in vertebrate physiology, particularly in regulating water and sodium ion balance. Upon activation, V2R engages Gαs and Gαq/11 signalling pathways, leading to subsequent recruitment of β-arrestins and internalisation of the receptor into endosomes.

Vasopressin exerts multifaceted effects on cardiovascular homeostasis, modulating water resorption in the kidneys, regulating smooth muscle tone in the vasculature, and serving as a central neurotransmitter involved in brainstem autonomic function [7]. Meanwhile, cAMP functions as a critical second messenger, orchestrating cellular metabolism and function in response to various extracellular hormones and neurotransmitters.

1.2 Problem Statement

Conventionally, studies on G protein-coupled receptor (GPCR) mediated signalling have primarily focused on the concentration of the stimulus ligand and the resulting cellular response. However, this approach overlooks the crucial aspect of quantifying the amount of information conveyed by the GPCR receptor channel.

Previous investigations have relied on measuring the concentration of intracellular calcium ions (Ca2+) as a proxy for GPCR activation, particularly in response to acetylcholine (Ach) stimulation in HEK293 cells [6]. However, the reliability of Ca2+ measurements is compromised due to limitations in experimental techniques, especially when capturing the transient bursts of Ca2+ release.

Building upon this experimental framework, our study aims to assess the channel capacity of GPCR signalling using cyclic adenosine monophosphate (cAMP) response data obtained from HEK293 cells stimulated with vasopressin. By developing algorithms and conducting comprehensive data analysis, we seek to elucidate the information-carrying capacity of the vasopressin type 2 receptor (V2R) in HEK293 cells. This endeavour will contribute to a deeper understanding of GPCR signalling dynamics and pave the way for novel insights into cellular communication processes.

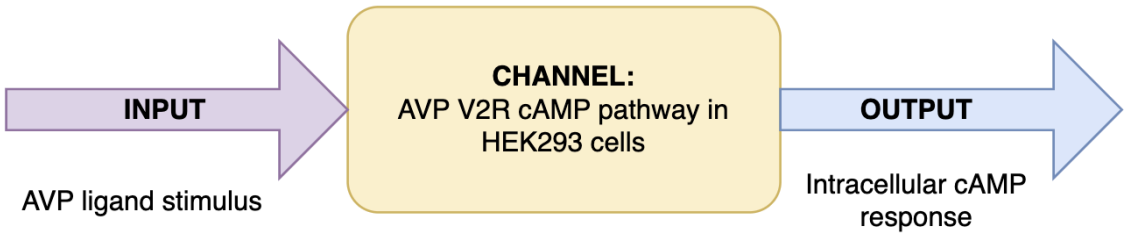

Figure 1a. Diagram illustrating the input,output and the channel within the biological system of HEK 293 cells, utilised for estimating channel capacity

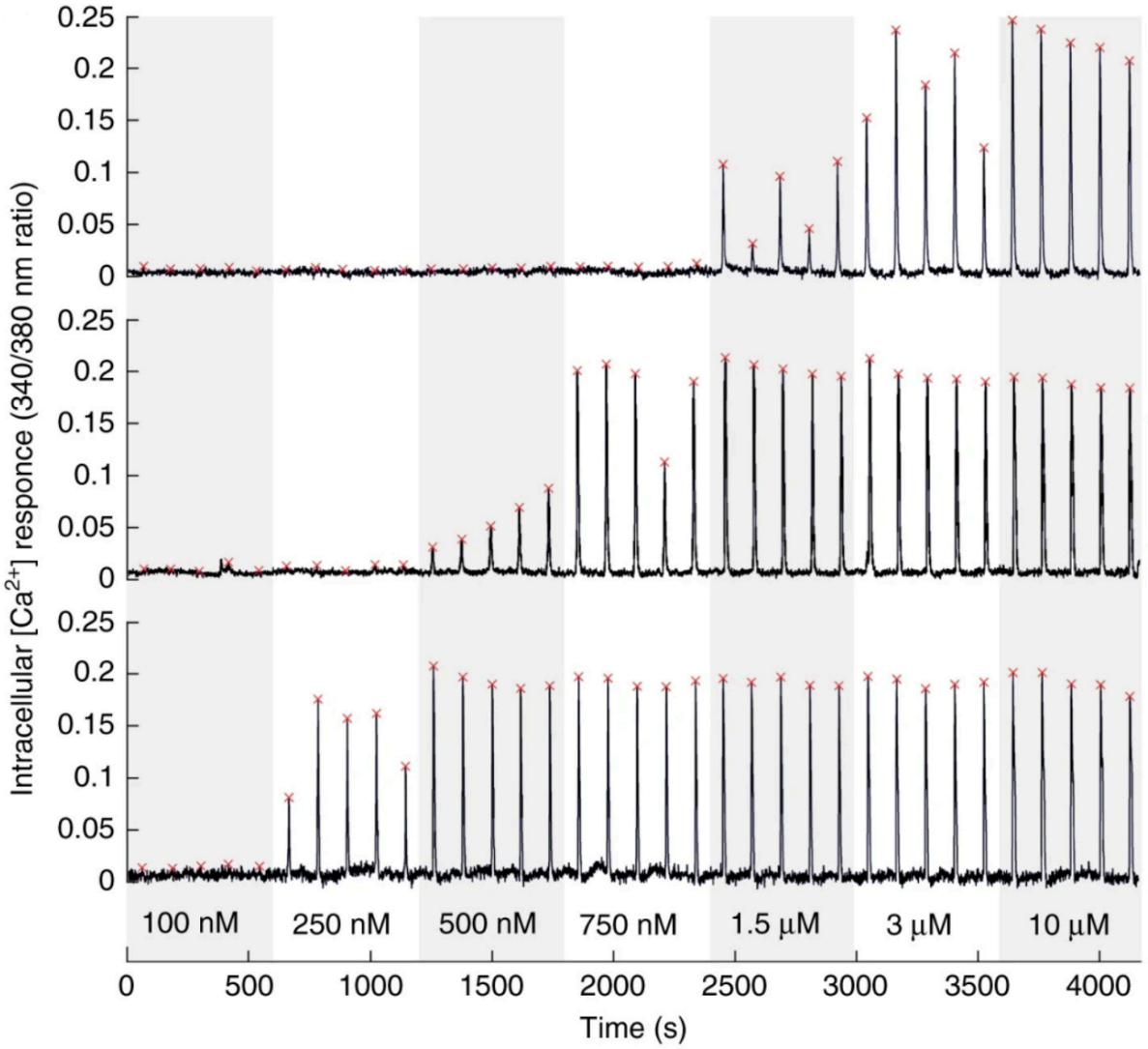

Figure 1b. Previous investigations have relied on measuring the concentration of intracellular calcium ions (Ca2+) as a proxy for GPCR activation, particularly in response to acetylcholine (Ach) stimulation in HEK293 cells.Three examples of different single cell's Ca2+ responses to repeated stimulations with increasing concentrations of Ach. It is evident that some cells respond only to high Ach concentrations (top example), while others are sensitive already to the low concentrations but rapidly reach the plateau in their response strength (bottom example). The example in the middle is an intermediate. [6]

1.3 Importance and Relevance of the Project

1.3.1 Implications of Information Theory in Cellular Signalling

The utilisation of information theory language in understanding cellular signalling unveils significant implications, particularly in recognizing certain disease states, such

as cancer, as 'information diseases'. During the transition from a healthy to a diseased state, cells undergo substantial changes in the information processing characteristics of specific signalling networks or pathways. Notably, the oncogenicity of G protein- coupled receptor (GPCR) signalling has been extensively evidenced, with GPCRs and their associated pathways emerging as prime targets for anti-cancer drugs and drug discovery efforts [10].

1.3.2 Distinguishing Cancerous from Non-cancerous Cells

Despite the acknowledged involvement of GPCRs in oncogenic processes, the application of information theory to dissect oncogenic cellular signalling remains largely unexplored. This gap persists partly due to the technical challenges associated with applying channel capacity assessment setups based on individual cell responses within a heterogeneous population to cancer cells [8]. The inherent cell heterogeneity in cancer cell biology complicates the analysis of signalling pathways, posing a significant obstacle in distinguishing between cancerous and non-cancerous cells.

1.3.3 Prevalence of GPCR Targeted Drugs

It is noteworthy that a considerable proportion of drugs available in the market target GPCRs or their associated signalling pathways. Notable examples include beta- blockers targeting beta-adrenergic receptors for cardiovascular diseases, antipsychotics targeting dopamine receptors for psychiatric disorders, and antihistamines targeting histamine receptors for allergies[9]. This prevalence underscores the critical importance of deepening our understanding of GPCR pathways, both in normal physiology and disease states, such as cancer.

2. Methods and Techniques

2.1 Cell Population, Stimulation and Experimental Replication

In this investigation, the cAMP (cyclic adenosine monophosphate) response of a cell ensemble comprising approximately 65,000 HEK 293 cells, stimulated with the AVP (arginine vasopressin) ligand, was scrutinised. The cellular population underwent evaluation across an array of 12 distinct ligand concentrations.

Each measurement iteration was repeated twice to ensure the robustness and reliability of the acquired data. The temporal dynamics of cAMP response were meticulously recorded, capturing variations across different time points for each concentration. Notably, peak responses were discerned for each concentration and repetition, serving as pivotal indicators of cellular response.

2.2 Data Processing and Normalisation

Following data acquisition, a logarithmic transformation of the cAMP responses was performed, facilitating a normalised representation amenable to comparative analyses across concentrations. Subsequently, the average and standard deviation of these transformed responses were calculated, providing key statistical descriptors for subsequent investigations.

Drawing upon principles rooted in information theory, particularly the framework of mutual information, the interrelationship between the input ligand concentration (designated as c) and the logarithmically transformed cAMP response (designated as r) was rigorously examined. Mutual information, symbolised as I(r,c), emerged as a pivotal metric elucidating the extent of information shared between these variables.

2.3 Characterization of Output Distribution and Channel Capacity Estimation

Central to our investigation was the exploration of channel capacity, delineating the upper limit of mutual information across all conceivable input distributions (P(c)). Paramount to this endeavour was the characterization of the output variable's distribution conditioned on the input variable, denoted as P(r | c). Empirical observations revealed a propensity for P(r | c) to adhere to a normal distribution, underscoring the underlying dynamics governing information transmission within the cellular milieu.
Moreover, parallel investigations were conducted at the single-cell level, wherein similar experiments were performed on multiple individual cells. The average channel capacity for a cell was calculated using the results obtained from these multiple single-cell experiments.

Notably, the analytical computations and statistical analyses were facilitated using a Python script tailored for this purpose, ensuring accuracy and efficiency in data processing and interpretation [Supplementary Material].

This comprehensive methodological framework enabled a thorough exploration of the efficiency of information transmission between ligand concentration and cAMP response, thereby offering nuanced insights into the intricate signalling dynamics orchestrating cellular responsiveness within the HEK 293 cellular context.

3. Data Collection, Analysis and Interpretation

Cellular signalling, as measured using chemical and biological tool sets although reflects the trend and toner of cellular signalling, not only that it also able to fractionate downstream signalling transcoders and effectors, but does not absolutely quantitate the absolute magnitude of signal amplification following external triggers, for simplicity can be called as the agonist. Here we designed an improved model to measure the channel capacity of GPCR as they work as efficient signalling conduits. Using an information theory-based model that is robustly used in electrical and computational networking studies, we assume GPCR signalling can also be quantitated accurately using this model. We here take a model GPCR i.e. vasopressin receptor (V2R), a well-studied prototypical receptor. We utilise previously published data that measures cAMP signalling in V2R [12]. The data set is chosen as it utilises cAMP signalling data not only from wild-type V2R but also deficient V2R mutant (V2R_360). It also entails Intrabody based signalling data that shows signalling rescue of the V2R_360 [Table 4] . Overall this data has robust signalling variation in the same receptor and provides a unique opportunity to test the validity of our improved model for measuring channel capacity. Subsequently, we also utilise a single-cell V2R signalling data set just to see how our model gage the improved experimental platform to study a single-cell homogeneous reaction platform [Figure 2]. Anticipating our model can differentiate between the diverse signalling landscape we will subsequently be encouraged to use this model to apply on a platform to differentiate between cancer and a normal cell.

Concentration	-6	-7	-8	-9	-10	-11	-12	-13	-14
Replicate 1	22536	23300	22054	9658	1234	366	246	246	266
Replicate 2	22890	24571	22124	10290	1298	366	266	268	256
Replicate 3	15777	17005	14963	4728	744	312	308	314	284
Replicate 4	16565	18347	18171	5372	804	374	336	360	328
Replicate 5	14655	15903	16745	16259	7788	648	184	172	158
Replicate 6	16003	17315	17895	17821	8608	798	216	208	184

Table 1. Heat map representation of peak cAMP response vs $\log_{10}$ AVP concentration for V2R_360 +IB Cntrl

Concentration	-6	-7	-8	-9	-10	-11	-12	-13	-14
Replicate 1	20004	21148	20992	9730	1426	388	258	282	278
Replicate 2	21290	20862	20084	9852	1634	376	276	304	286
Replicate 3	22280	25241	23838	8424	1170	484	462	494	514
Replicate 4	24140	24026	23784	8650	1464	542	454	458	474
Replicate 5	21360	24276	25479	24288	11980	1030	278	256	224
Replicate 6	21386	24050	25343	23818	10964	948	310	252	248

Table 2. Heat map representation of peak cAMP response vs $\log_{10}$ AVP concentration for V2R_WT+IB 30

Concentration	-6	-7	-8	-9	-10	-11	-12	-13	-14
Replicate 1	19233	19939	18577	10164	1434	338	270	284	284
Replicate 2	19979	20608	19071	10538	1446	340	268	290	272
Replicate 3	15681	17203	17505	7070	1260	466	386	412	370
Replicate 4	15759	18277	17325	7572	1230	436	372	370	306
Replicate 5	15555	16283	16653	15989	8910	750	218	174	184
Replicate 6	15697	16071	16137	15391	8486	762	200	192	162

Table 3. Heat map representation of peak cAMP response vs log_{10} AVP concentration for V2R_WT+ IB Cntrl

Concentration	-6	-7	-8	-9	-10	-11	-12	-13	-14
Replicate 1	23330	24010	23492	13544	1898	466	388	370	358
Replicate 2	23870	24759	24551	13584	1828	450	348	350	342
Replicate 3	22944	24404	24322	10452	1646	618	564	520	514
Replicate 4	20516	23984	22470	9440	1462	548	488	472	418
Replicate 5	20308	22538	23654	21890	11516	966	286	260	218
Replicate 6	19119	21614	22858	20760	12652	950	280	254	230

Table 4. Heat map representation of peak cAMP response vs log_{10} AVP concentration for V2R_360 + IB 30

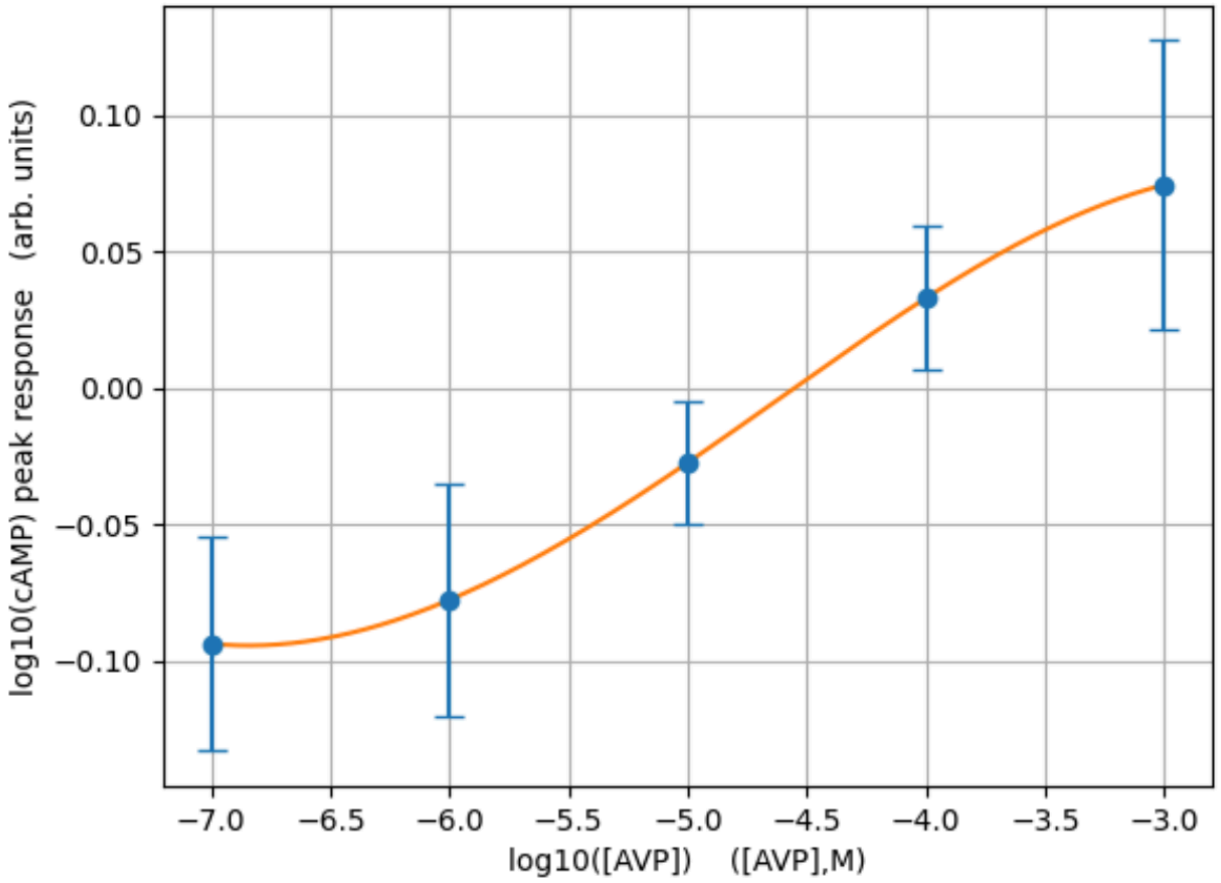

Figure 2. Best fit regression curve of log_{10}cAMP response vs log_{10}AVP concentration for individual HEK293 cell (n=3 replicates)

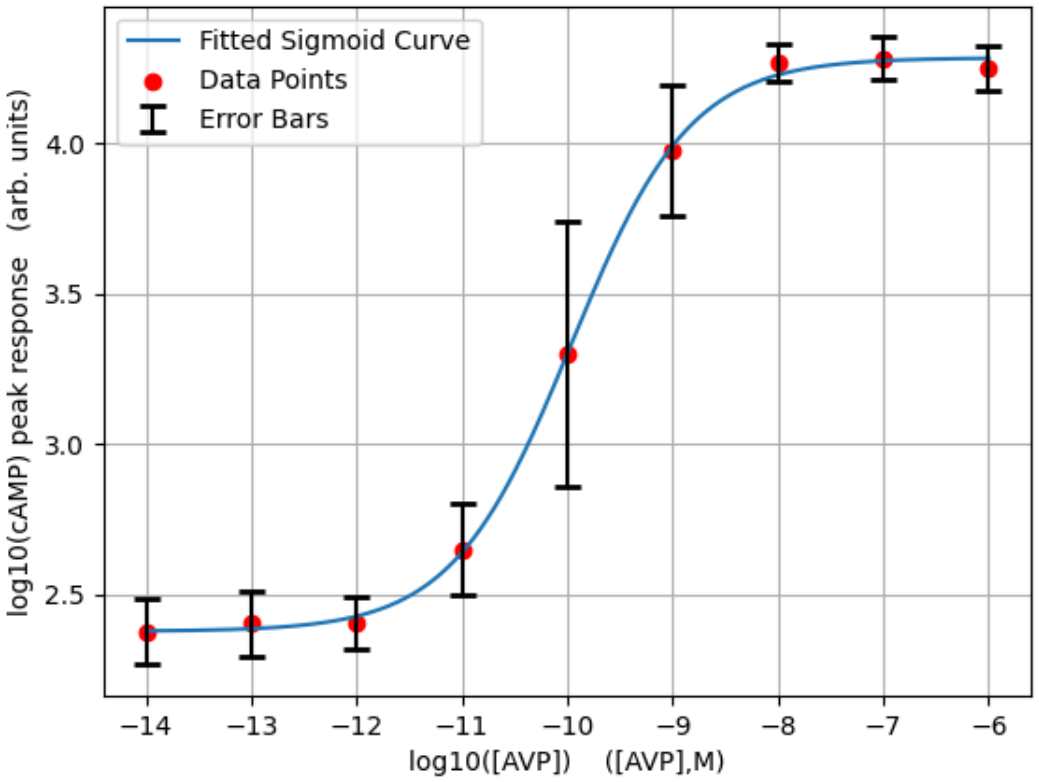

Figure 3. Best fit regression curve of log_{10}cAMP response vs log_{10}AVP in an ensemble of ~65000 HEK293 cells (n=6 replicates) V2R_WT

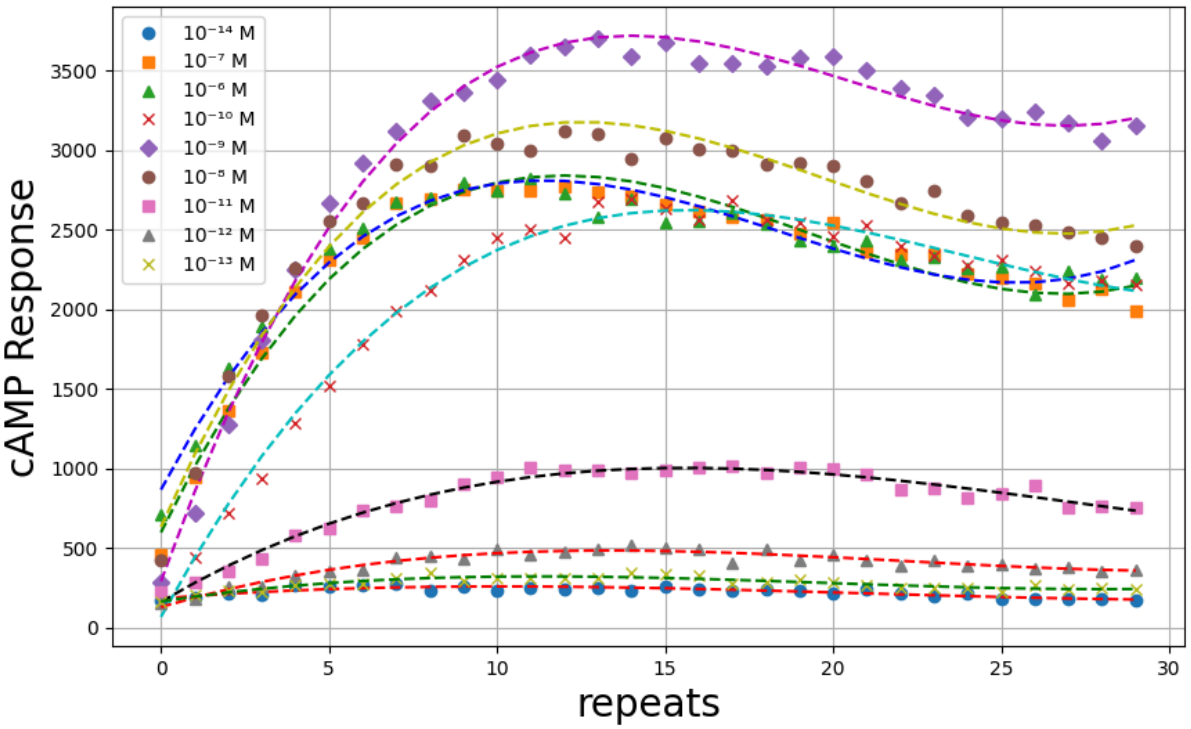

Figure 4. Temporal dynamics of cAMP response in HEK293 cell ensemble under varied AVP concentration stimulation

5. Channel Capacity Calculation

The channel capacity is the maximal rate of information that can be transmitted through a memoryless channel. This concept was developed within the framework of information theory initiated by Shannon [3][4][5][6].
More precisely the channel capacity K is given by max I(r, c), where I is the mutual information between the channel input c and the channel output r. In our setup, c is given by the AVP concentration, while r is the measured intracellular cAMP response.
The mutual information between c and r has been shown to be equal to

$$\int\int P(r,c)log(\frac{P(r,c)}{P(r)(P(c)})drdc$$

where $P(r,c)$ is the joint distribution of c and r
Since $P(r,c) = P(r|c)P(c)$, we have

$$I = \int\int P(r|c)P(c)log(\frac{P(r,c)}{P(r)(P(c)})drdc$$

Moreover, by marginalising over c, we have that

$$P(r) = \int P(r|c)P(c)dc, \text{ so that}$$

$$I = \int\int [P(r|c)P(c)(log(P(r|c) - log(\int P(r|c)P(c)dc))]drdc$$

A standard optimization procedure was used to find the discrete distribution over the cj that maximises I (under the constraint that $\sum_j P(cj) = 1$).

In so doing, we only estimate a lower bound for the capacity, because when maximising I over the input distribution, we can only consider a subset of distributions, namely the ones that are zero for all $c \neq cj$ $(j = 1, 2, 3 \ldots n)$.
This means that we are maximising over a subset of distributions, necessarily resulting in a lower bound estimate of the channel capacities.
The python script and data used to estimate the channel capacity can be accessed at *https://github.com/TsarIM/Estimating-GPCR-Signalling-potential-Modelling-to-Quantitate-Channeling-Capacity-with-cAMP-Data.git*

6. Results

Ca^{2+} measurements used in previous literature are deemed unreliable due to their susceptibility to variability and inconsistency, particularly because Ca^{2+} release occurs in bursts and can vary across different laboratory settings [6][Figure 1b]. Our study circumvented this issue by utilising cAMP data to calculate channel capacity. Using the above-cited data, we measured the channel capacity of V2 receptor, as it is one of the most studied prototypical GPCR, using our improved model as discussed in the previous section. We observed that the V2 receptor wild type exhibits normal signalling, with similar values observed in both intrabody control and 30 conditions [Table 5]. Consistent with previous literature , V2 360 shows a reduced signalling capacity, resulting in a dip in channel capacity. Intrabody 30 appears to rescue the channel capacity of the V2 360 receptor, yielding a similar value of 2.28336056 bits [12]. However, in the presence of the IB 30 intrabody, receptor functionality is restored to wild-type levels, thereby augmenting information-carrying capacity [Table 5].

Next, channel capacity was calculated using focused and homogeneous measurements of individual cells, revealing a higher estimate of 2.364370355 bits. Our application of the model for channel capacity to single-cell resolution demonstrated that focused cell studies resulted in a higher channel capacity compared to cell ensemble measurements [Figure 2].

Sample	**Channel Capacity (bits)**
V2R_WT + IB cntrl	2.242668029
V2R_WT + IB 30	2.166057747
V2R_360 + IB cntrl	1.940012271
V2R_360 + IB 30	2.28336056
Single cell	2.364370355

Table 5. Channel Capacity estimates for different cell samples

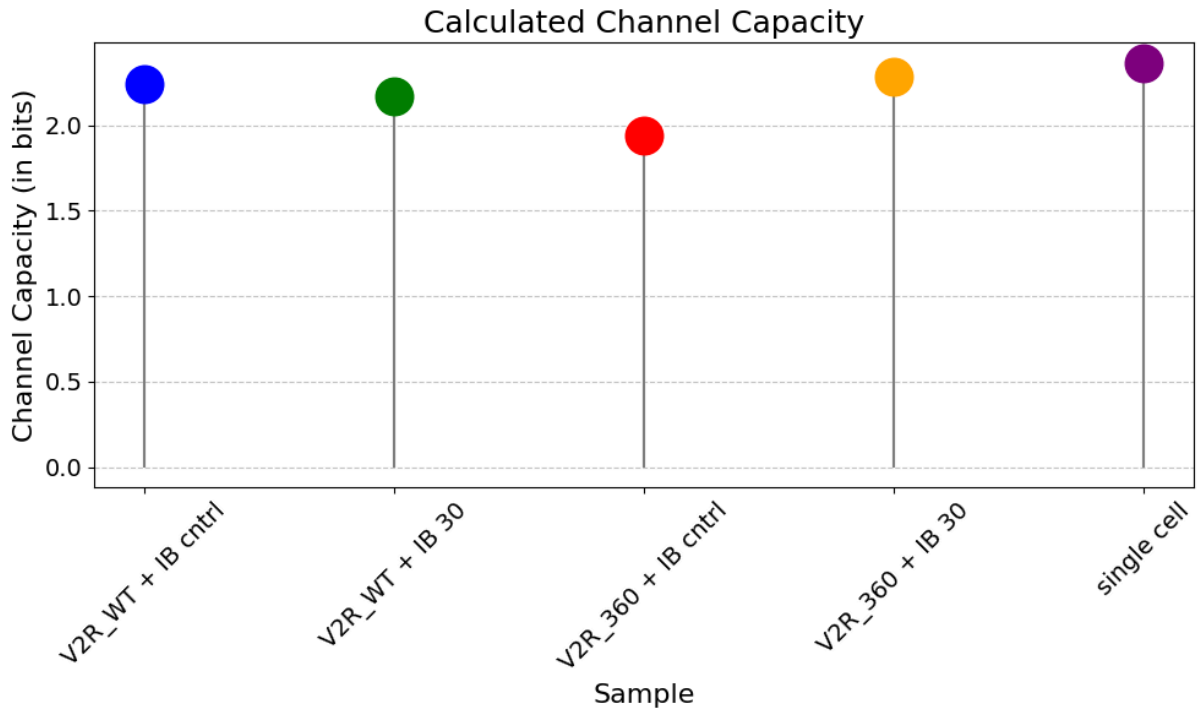

Figure 5. Channel Capacity estimates for different cell samples

7. Discussion

7.1 Interpretation

Our findings reveal that a single cell possesses high-capacity information transmission channels, approximately 2.36 bits, enabling it to discern between $2^{2.36}$ different concentrations. This capability allows cells to effectively respond to varying levels of external stimuli. Individual cells within a population exhibit dynamic ranges adapted to different absolute agonist levels. These conclusions provide insights into well-established processes in cell and developmental biology, such as tissue morphogenesis, where cells acquire distinct fates based on the concentration of received morphogens [11].

To test the utility of the model we compared the V2R_WT receptor with a signalling deficient mutant V2R_360 in the presence and absence of signalling recusing intrabody IB 30. A comparison between V2R_360 and V2R_WT in the presence of IB Ctl (control Intrabody) reveals a decrease in channel capacity in V2R_360, suggesting functional impairment (Figure 5 and Table 5). However, in the presence of the IB 30 intrabody receptor functionality is rescued and restored to wild-type levels, augmenting information-carrying capacity. The data not only assess the potential signalling rewiring capacity of designer intrabody and their therapeutic implication but also reaffirms the utility of our model that maps the variation of channel capacity landscape that was induced in the experimental framework. This trend in channel capacity variation is in line with the observed outcome in previous works [12], suggesting the model is robust

and could be explored for detecting cancer cells, where disruption of signaling is a *prima facie* event.

We went to apply our model for channel capacity to single-cell resolution, where focused cell study resulted in higher channel capacity as compared to cell ensemble. The finding that individual cell channel capacity is higher than that of that calculated in cell ensembles may well be confounded by the fact that different cells respond with different strengths to the same concentration of the stimulus. Thus, measurements averaging over many cells can effectively result in an estimate of the channel capacity that is much smaller than the intrinsic capacity of a single cell [13].

7.2 Comparison with existing literature

In contrast to Ca^{2+} response examined in existing literature, our research focused on the cAMP response in HEK293 cells stimulated with vasopressin. From our investigation, we calculated a channel capacity value of approximately 2.16± 0.13 bits (mean± sd) for the cell populations under consideration and 2.36 bits for the single cell. This result indicates a similar information transmission efficiency of AVP V2R GPCR compared to the Ach M3R GPCR response studied in the existing literature. Specifically, for the n = 433 cells analysed using a parametric distribution approach, the average channel capacity is reported at 2.06 ± 0.31 bits (mean ± sd). Additionally, without parameter interpolation, a lower bound estimate of 1.65 bits was reported [6].

7.3 Analysis of limitations and potential sources of error

In our investigation of quantifying the information capacity of GPCR signalling using cAMP response data in HEK293 cells stimulated with vasopressin, several limitations and potential sources of error were identified. These include variability in experimental design, data quality concerns, reliance on statistical assumptions, potential oversimplification of biological processes in mathematical models, limited generalizability to diverse cellular contexts, and the risk of interpretation bias. Addressing these challenges through optimised experimental protocols, rigorous quality control measures, sensitivity analyses, more sophisticated modelling techniques, cautious extrapolation of findings, and ensuring objectivity in interpretation is essential for enhancing the reliability and robustness of our research findings and advancing our understanding of GPCR signalling dynamics.

References

1. Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol. 2003 Jun;63(6):1256-72.
2. Nambi, P. & Aiyar, N. G protein-coupled receptors in drug discovery. Assay Drug Dev. Technol. 1, 305–310 (2003).
3. Cover, T. M. & Thomas, J. A. Elements of Information Theory 2nd edn (John Wiley & Sons, Inc., 2006), Hoboken, New Jersey.
4. Shannon, C. E. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423 (1948).
5. Shannon, C. E. A mathematical theory of communication. Bell Syst. Tech. J. 27, 623–656 (1948).
6. Keshelava, A., Solis, G.P., Hersch, M. et al. High capacity in G protein-coupled receptor signalling. Nat Commun 9, 876 (2018).
7. Holmes, C.L., Landry, D.W. & Granton, J.T. Science Review: Vasopressin and the cardiovascular system part 1 – receptor physiology. Crit Care 7, 427 (2003).
8. Koval A, Zhang X, Katanaev VL. Improved approaches to channel capacity estimation discover compromised GPCR signalling in diverse cancer cells. iScience. 2023 Jul 3;26(8):107270.
9. Sriram K, Insel PA. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? Mol Pharmacol. 2018 Apr;93(4):251-258.
10. Nieto, Ainhoa & McDonald, Patricia. (2017). GPCRs: Emerging anti-cancer drug targets. Cellular Signalling. 41. 10.1016/j.cellsig.2017.09.005.
11. Gilbert, S. F. Developmental Biology 10th edn (Sinauer Associates, 2014), Sunderland, Massachusetts.
12. Mithu Baidya, M., Chaturvedi, M., Dwivedi-Agnihotri, H. et al. Allosteric modulation of GPCR-induced β-arrestin trafficking and signalling by a synthetic intrabody. Nat Commun 13, 4634 (2022).
13. Bao, X. R., Fraser, I. D., Wall, E. A., Quake, S. R. & Simon, M. I. Variability in G-protein-coupled signalling studied with microfluidic devices. Biophys. J. 99, 2414–2422 (2010).

List of Images

List Of Tables

Supplementary Material

Code and data supporting the findings of this manuscript can be accessed at *https://github.com/TsarIM/Estimating-GPCR-Signalling-potential-Modelling-to-Quantitate-Channeling-Capacity-with-cAMP-Data.git*