


 Cite this: *Chem. Commun.*, 2021, 57, 12305

 Received 16th June 2021,
 Accepted 12th October 2021

DOI: 10.1039/d1cc03202j

rsc.li/chemcomm

Fragment-based design of selective GPCR ligands guided by free energy simulations†

 Pierre Matricon,^a Duc Duy Vo,^a Zhan-Guo Gao,^b Jan Kihlberg,^c
 Kenneth A. Jacobson^b and Jens Carlsson^{b,*a}

Fragment-based drug discovery relies on successful optimization of weakly binding ligands for affinity and selectivity. Herein, we explored strategies for structure-based evolution of fragments binding to a G protein-coupled receptor. Molecular dynamics simulations combined with rigorous free energy calculations guided synthesis of nanomolar ligands with up to >1000-fold improvements of binding affinity and close to 40-fold subtype selectivity.

Technologies that enable more efficient generation of lead candidates are needed to increase the success rate of drug discovery. The small fraction of chemical space that can be screened experimentally continues to be a major limitation. Even if high-throughput screening (HTS) and DNA-encoded libraries contain millions to billions of compounds, these only explore a small fraction of the 10^{60} possible drug-like molecules and will lack chemical starting points for many targets.^{1,2} Fragment-based drug discovery (FBDD) takes an alternative route to identify leads, which has already led to several clinical candidates.³ By first screening compounds that are less than half the size of a drug, fragment libraries can provide better coverage of chemical space.² Another advantage is that fragments are likelier to bind to a protein than drug-like compounds because of their small size and low molecular complexity.⁴ However, the high hit rates from fragment screening comes at a price – the compounds will bind weakly and not be selective for the target.⁵ The second step of FBDD, fragment-to-lead optimization, can be very challenging, in particular if a crystal structure of the protein–fragment complex is not

available. The dependence of FBDD on high resolution structures has limited the applicability of the method for important drug targets such as transmembrane receptors.⁶ For these reasons, accurate computational models of fragment binding and methods to guide optimization would be valuable. Several recent studies suggest that relative binding free energies calculated from molecular dynamics (MD) simulations can guide hit-to-lead optimization for important drug targets such as G protein-coupled receptors (GPCRs).^{7–9}

In this work, we undertook the challenge to optimize fragments binding to a GPCR with the goal to investigate three central questions in FBDD. Firstly, can atomic resolution GPCR structures guide fragment optimization? Crystal structures of numerous GPCR drug targets have recently been solved,¹⁰ but as complexes with fragments remain scarce, we used MD simulations to model receptor–fragment interactions. As there are multiple GPCR subtypes recognizing the same ligand, it was essential to achieve both affinity and selectivity. Our second question was if MD simulations can be used to model how binding affinity and selectivity is affected by small changes to a fragment's chemical structure. We assessed if rigorous free energy methods could predict the affinities of evolved fragments. Finally, we analysed advantages of using FBDD to develop chemical probes. A prospective study was performed by iteratively designing elaborated fragments based on the receptor structures and performing free energy calculations. Compounds were synthesized and tested in pharmacological assays, followed by analysis of the accuracy of the computational predictions.

Recently determined crystal structures of the A₁ and A_{2A} adenosine receptors (A₁- and A_{2A}ARs) could facilitate development of drugs to treat cancer, CNS, and cardiovascular diseases.^{11,12} Prior to the release of the A₁AR structure, we discovered a fragment (benzothiazole 1) binding to this GPCR with an affinity of 11.2 μM (Fig. 1).⁹ Docking of the fragment to the A₁AR binding site suggested that it was anchored by hydrogen bond interactions with Asn254. The ligand interacted with residues that are conserved in both receptors, but the structures revealed a residue substitution that created a unique hydrophobic subpocket in the A₁AR close to the amide moiety

^a Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, Uppsala SE-751 24, Sweden. E-mail: jens.carlsson@icm.uu.se

^b Molecular Recognition Section, Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. E-mail: kennethj@nidk.nih.gov

^c Department of Chemistry – BMC, Uppsala University, Uppsala SE-751 23, Sweden

† Electronic supplementary information (ESI) available: Methods and supplementary results for computational chemistry, compound synthesis, and biological assays. See DOI: 10.1039/d1cc03202j



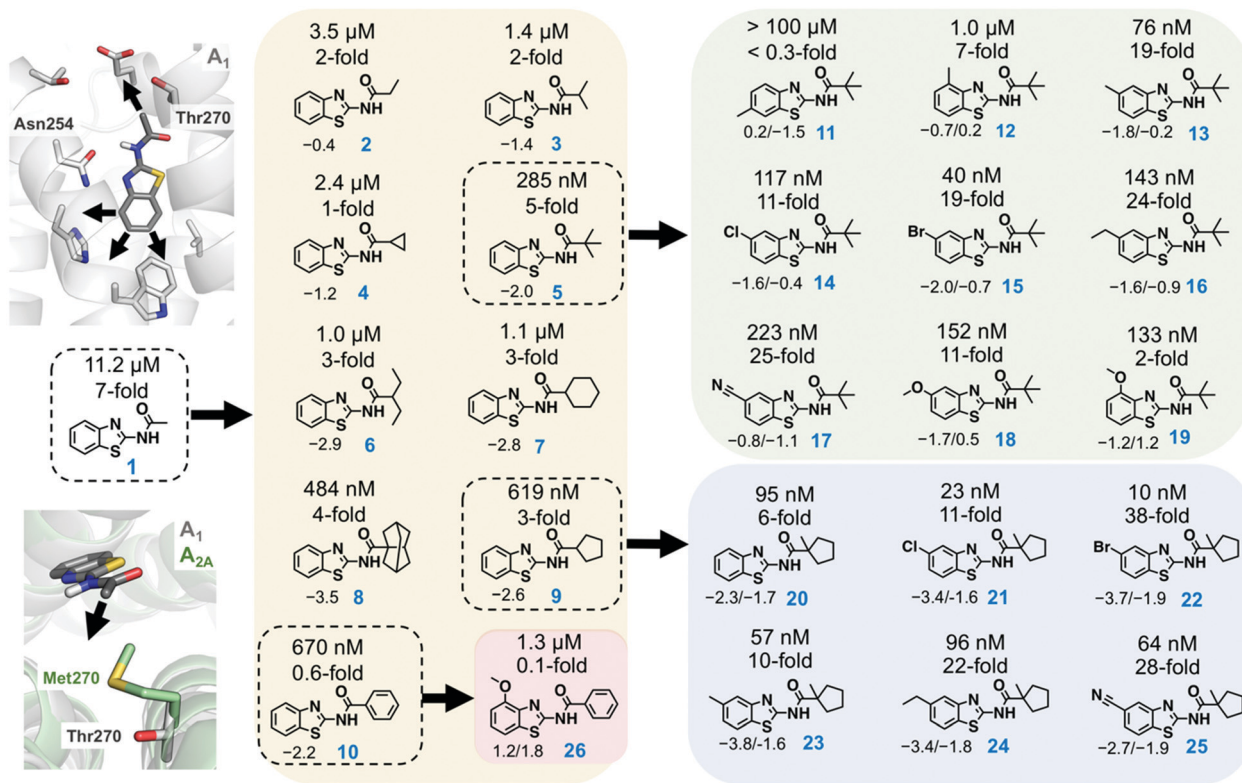


Fig. 1 Fragment-based lead design guided by free energy simulations. Growth vectors to evolve compound **1** were identified in the A₁AR binding site. A first series of compounds explored a unique subpocket of the A₁AR at the extracellular entrance of the binding site (**2–10**). Two additional series of compounds resulted in high affinity and subtype selective ligands (**11–26**). The experimentally determined binding affinities (K_i values) and selectivity for the A_{2A}AR are shown above each compound. The compounds used as references for the FEP calculations are marked with black dashed lines. For compounds **2–10**, the calculated relative binding free energies (kcal mol⁻¹) at the A₁AR are shown below each compound and negative values indicate improved affinity. For compounds **11–26**, the relative binding free energies at the A₁AR and the difference in relative binding free energy between A₁AR and A_{2A}AR (negative values indicate improved A₁AR selectivity) are shown below each compound (affinity/selectivity, kcal mol⁻¹). The experimental and calculated binding data are also summarized in Tables S1–S4 (ESI[†]).

of the fragment. This pocket is located the extracellular entrance of the binding site and was not accessible in the A_{2A}AR because Thr270 is replaced by the bulkier side chain of Met270 (Fig. 1). The fragment only had 13 heavy atoms (HAs), corresponding to a ligand efficiency (LE, binding free energy per heavy atom) of 0.52 kcal mol⁻¹ HA⁻¹, which would be considered to be an excellent starting point for FBDD.¹³

The pocket identified in the A₁AR crystal structure guided optimization of the fragment. Compounds **2–10** explored a single growth vector by positioning hydrophobic substituents of varying size and shape in the pocket (Fig. 1). The calculations were performed to predict binding free energies of these compounds relative to **1**. MD simulations were initiated from an A₁AR crystal structure,¹¹ which was equilibrated in the presence of a lipid bilayer and water. In the simulations, predicted ligands were alchemically transformed into **1** in complex with the A₁AR and in aqueous solution using the free energy perturbation (FEP) technique, which can be used to calculate relative binding affinities based on a thermodynamic cycle.⁹ The FEP calculations were carried out using the program Q,¹⁴ and an average simulation length of >0.4 μs was used for each compound pair. Detailed simulation protocols are described in the ESI[†]. FEP predicted that all the designed compounds would show improved A₁AR affinity, but there was a large

variation in the magnitude of the gain of binding free energy (Fig. 1 and Tables S1, S2, ESI[†]). Subsequently, **2–10** were synthesized, and evaluated in radioligand binding assays at the A₁AR. Detailed synthesis and assay procedures are available in the ESI[†]. As predicted by FEP, all compounds had improved affinity compared to **1** (Fig. 1). The three compounds with the smallest substituents (**2–4**) displayed the smallest gains of binding, which agreed with the FEP predictions. However, the computational ranking of the larger substituents (**5–10**) did not correlate well with the experimental data. Except in the case of compound **5**, the free energy gain was overestimated by the simulations. Pivaloylated **5** had the highest affinity ($K_i = 285$ nM), corresponding to a 39-fold improvement, and the LE increased from 0.52 to 0.56 kcal mol⁻¹ HA⁻¹. Compound **9** ($K_i = 619$ nM) was slightly weaker than **5**, but also retained a high LE (0.50 kcal mol⁻¹ HA⁻¹).

Based on the first set of compounds, **5** and **9** were further elaborated. Interestingly, binding data (Table S1, ESI[†]) showed that these ligands were not more selective for the A₁AR compared to **1** despite that both positioned substituents in the non-conserved pocket (Fig. 1). MD simulations of the complexes with the A₁AR and A_{2A}AR indicated that this result was due to the small size of these ligands. The ligands were able to adopt slightly different binding modes in the receptors,



which reduced clashes with Met270 in the A_{2A} AR. In a second step, we explored other growth vectors in the binding site to improve selectivity (Fig. 1), which involved adding substituents at either the 4-, 5-, or 6-position of the benzothiazole moiety. The rationale behind these designs was that improved anchoring of the fragments would rigidify the binding mode and thereby enhance the effect of having substituents in the non-conserved pocket, potentially leading to selectivity. As 5 had the highest affinity, the first series of elaborations focused on optimizing this fragment. The FEP protocol was used to calculate binding free energies relative to 5 in the A_1 AR and A_{2A} AR binding sites to predict changes in affinity as well as selectivity. A methyl substituent was first introduced at three different positions (compound 11–13) of the benzothiazole moiety to probe the possible growth vectors (Fig. 1). The simulation results indicated that the 5-position was a hotspot for increasing A_1 AR affinity. As FEP also indicated that selectivity could be improved, the synthesis (see ESI† for details) was focused on substituents at this position (Fig. 1). The experimental binding data showed that FEP correctly predicted that the 5-position (13) led to the largest improvement of A_1 AR affinity among the three methyl substituents (Fig. 1 and Tables S3, S4, ESI†). The large loss of affinity for 11 was not captured by FEP, but the ranking of the three compounds by affinity was correct. Several of the substitutions at the 5-position (13–18) resulted in improved A_1 AR selectivity and affinity, which was predicted by FEP in a majority of the cases. The lack of improvement of A_1 AR selectivity for 19, which had a 4-methoxy substituent, was also in agreement with the calculated free energies. Compound 15, which had a 5-bromo substituent, showed the highest A_1 AR affinity (40 nM), was 19-fold selective, and had a remarkable LE of 0.60 kcal mol⁻¹ HA⁻¹.

Compound 9 was optimized in two steps and involved synthesis and experimental evaluation of six additional compounds (see ESI† for details, Table S3, ESI†). As 9 had weaker affinity than 5 at the A_1 AR, a methyl group was added to the cyclopentyl moiety to mimic the *tert*-butyl group of 5. The resulting compound 20 was only 6-fold selective for the A_1 AR, but its experimental affinity improved to 95 nM (Fig. 1). Additional FEP calculations of relative affinities and selectivity were performed for substituents at the 5-position, which led to synthesis of 21–25. Compound 22, which had a 5-bromo substituent, resulted in the highest affinity (K_i = 10 nM) and a 38-fold selectivity for the A_1 AR (Fig. 2). FEP predicted that 22 would have higher affinity than 20, but only indicated a slight increase of selectivity (Fig. 1 and Tables S3, S4, ESI†). The overall improvement of 22 was astonishing. Addition of six heavy atoms to compound 1 resulted in >1000-fold increase of affinity, and selectivity was improved from 7- to 38-fold. MD simulation snapshots provided an explanation of the high affinity and selectivity of 22. The bulky 1-methylcyclopentyl substituent of 22 led to clashes with Met270 in the A_{2A} AR, which pushed the compound towards the extracellular loops. In contrast, this substituent fitted very well in the A_1 AR pocket with the 5-bromo substituent buried deeply in the binding site (Fig. 3).

In a final set of simulations, we explored if selectivity could instead be shifted towards the A_{2A} AR subtype. In fact,

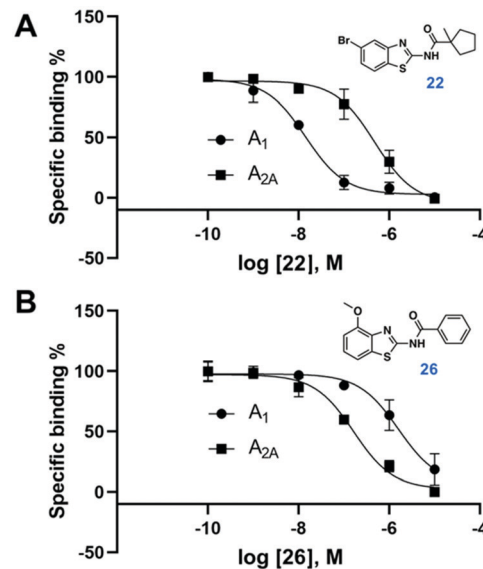


Fig. 2 Dose–response curves from radioligand binding assays for compounds 22 and 26 ($n = 2-4$).

benzothiazoles have generally been described as an A_{2A} AR selective scaffold.¹⁵ Compound 10 showed weak 2-fold A_{2A} AR selectivity. To test if selectivity could be further increased for the A_{2A} AR, compound 26, which had a 4-methoxy group on the benzothiazole ring, was synthesized (see ESI† for details). FEP predicted this substituent to shift affinity and selectivity further towards A_{2A} AR, which was also confirmed experimentally (106 nM, 13-fold selectivity, Fig. 2 and Tables S3, S4, ESI†).

Functional assays measuring intracellular concentrations of cAMP were performed at the human A_1 AR and A_{2A} AR to determine the efficacy of compounds 22 and 26. The experiments demonstrated that the compounds acted as antagonists of both receptors and the selectivity profiles were the same as in the binding assays (Fig. S1 and S2, ESI†). These results also agreed with the fact that the simulations were performed using an inactive receptor conformation.

To evaluate the overall performance of the FEP calculations in the affinity and selectivity optimization, the correlation between predicted and experimental binding free energies were

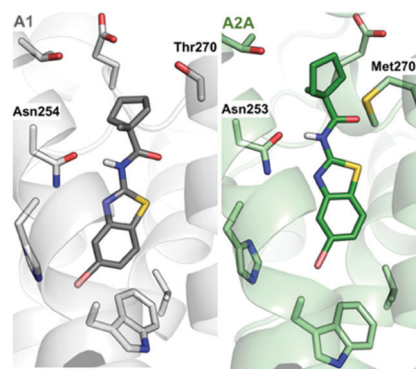


Fig. 3 Predicted structures of compound 22 bound to the A_1 - (grey) and A_{2A} AR (green).



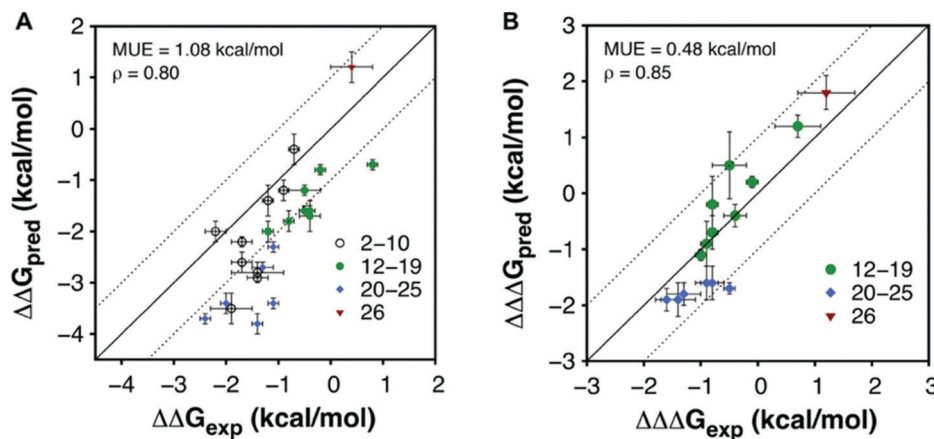


Fig. 4 Correlation between experimental and calculated free energies (A) for affinity and (B) selectivity at the A₁AR.

evaluated (Fig. 4 and Tables S2, S4, ESI[†]). For the affinity optimization at A₁AR the relative binding free energies were predicted with a mean unsigned error (MUE) of 1.08 kcal mol⁻¹, and there was a strong Spearman rank correlation ($\rho = 0.80$) for the 24 compounds with a determined affinity. Predictions of selectivity were evaluated based on the difference between relative binding free energies for the two receptors from FEP and experiment, and in this case the MUE and ρ were 0.48 kcal mol⁻¹ and 0.85, respectively, for the compounds with determined affinities. Hence, there was a good correlation between FEP and experiment for both affinity and selectivity. Returning to the three questions that motivated this study, we first conclude that computational models of GPCR-fragment complexes successfully guided fragment elaboration. Structure-informed selection of substituents led to a remarkable >1000-fold improvement of affinity and close to 40-fold receptor subtype selectivity. Second, the relative binding free energies calculated with FEP accurately ranked compounds by affinity as well as selectivity, suggesting that MD simulations can be a useful tool to identify which fragments to elaborate and guide selection of substituents. The approach is most suitable for receptors with a well-defined binding site that recognize fragment-sized endogenous ligands (e.g. aminergic GPCRs). Applications to other targets (e.g. peptide or protein-binding GPCRs) may be more challenging because fragments bind weakly and modelling of binding modes will be difficult. Finally, comparison of **22** to previously developed A₁AR ligands clearly illustrates the benefits of using FBDD compared to HTS, which initiates optimization from larger lead-like compounds. In fact, analysis of reported A₁AR ligands of the same size as **22** (<20 HA) showed that this compound is one of the highest affinity ligands of this size ever discovered (Table S5, ESI[†]). By carefully evolving fragments, atom-by-atom, high affinity leads can be obtained by synthesizing a small series of compounds, and the chances of obtaining leads with favourable physicochemical properties increase.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement: 715052). The work was also supported by the Swedish Research Council (2017-4676) and the Swedish strategic research programme eSENCE. K. A.

J. thanks the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Intramural Research Program (ZIADK31117) for financial support. Computational resources were provided by the Swedish National Infrastructure for Computing (SNIC).

Conflicts of interest

There are no conflicts to declare.

Notes and references

- V. Blay, B. Tolani, S. P. Ho and M. R. Arkin, *Drug Discovery Today*, 2020, **25**, 1807–1821.
- J. L. Reymond, *Acc. Chem. Res.*, 2015, **48**, 722–730.
- D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke and H. Jhoti, *Nat. Rev. Drug Discovery*, 2016, **15**, 605–619.
- A. R. Leach and M. M. Hann, *Curr. Opin. Chem. Biol.*, 2011, **15**, 489–496.
- S. Schultes, C. De Graaf, E. E. J. Haaksma, I. J. P. De Esch, R. Leurs and O. Krämer, *Drug Discovery Today: Technol.*, 2010, **7**, e157–e162.
- S. P. Andrews, G. A. Brown and J. A. Christopher, *ChemMedChem*, 2014, **9**, 256–275.
- E. B. Lenselink, J. Louvel, A. F. Forti, J. P. D. Van Veldhoven, H. De Vries, T. Mulder-Krieger, F. M. McRobb, A. Negri, J. Goose, R. Abel, H. W. T. Van Vlijmen, L. Wang, E. Harder, W. Sherman, A. P. Ijzerman and T. Beuming, *ACS Omega*, 2016, **1**, 293–304.
- F. Deflorian, L. Pérez-Benito, E. B. Lenselink, M. Congreve, H. Van Vlijmen, J. S. Mason, C. de Graaf and G. Tresadern, *J. Chem. Inf. Model.*, 2020, **60**, 5563–5579.
- P. Matricou, A. Ranganathan, E. Warnick, Z. G. Gao, A. Rudling, C. Lambertucci, G. Marucci, A. Ezzati, M. Jaitheh, D. Dal Ben, K. A. Jacobson and J. Carlsson, *Sci. Rep.*, 2017, **7**, 6398.
- M. Congreve, C. de Graaf, N. A. Swain and C. G. Tate, *Cell*, 2020, **181**, 81–91.
- R. K. Y. Cheng, E. Segala, N. Robertson, F. Deflorian, A. S. Doré, J. C. Errey, C. Fiez-Vandal, F. H. Marshall and R. M. Cooke, *Structure*, 2017, **25**, 1275–1285.
- A. Glukhova, D. M. Thal, A. T. Nguyen, E. A. Vecchio, M. Jörg, P. J. Scammells, L. T. May, P. M. Sexton and A. Christopoulos, *Cell*, 2017, **168**, 867–877.
- A. L. Hopkins, G. M. Keserü, P. D. Leeson, D. C. Rees and C. H. Reynolds, *Nat. Rev. Drug Discovery*, 2014, **13**, 105–121.
- J. Marelus, K. Kolmodin, I. Feierberg and J. Åqvist, *J. Mol. Graphics Modell.*, 1998, **16**, 213–225.
- S. Basu, D. A. Barawkar, S. Thorat, Y. D. Shejul, M. Patel, M. Naykodi, V. Jain, Y. Salve, V. Prasad, S. Chaudhary, I. Ghosh, G. Bhat, A. Quraishi, H. Patil, S. Ansari, S. Menon, V. Unadkat, R. Thakare, M. S. Seervi, A. V. Meru, S. De, R. K. Bhamidipati, S. R. Rouduri, V. P. Palle, A. Chug and K. A. Mookhtiar, *J. Med. Chem.*, 2017, **60**, 681–694.

