Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

German Edition: DOI:

A Combination of Spin Diffusion Methods for the Determination of Protein–Ligand Complex Structural Ensembles**

*Jens Pilger, Adam Mazur, Peter Monecke, Herman Schreuder, Bettina Elshorst, Stefan Bartoschek, Thomas Langer, Alexander Schiffer, Isabelle Krimm, Melanie Wegstroth, Donghan Lee, Gerhard Hessler, K.-Ulrich Wendt, Stefan Becker, and Christian Griesinger**

anie_201500671_sm_miscellaneous_information.pdf

Supplement I

Experimental Section and Glycogen Phosphorylase Example

Experimental Section

Protein expression and purification: The chinese hamster Cα catalytic subunit of cyclic adenosine monophosphate (cAMP) dependent protein kinase A (PKA) was expressed and purified as published[1]. The development of a new expression and purification protocol of the human sEH hydrolase domain is described in detail in supplement II.

X-ray crystallography: Crystal structures of the PKA:L3 and PKA:L4 complexes were determined at 2.1 Å and 2.0 Å resolution using described purification and crystallization protocols and standard crystallographic procedures[2]. Crystal structures of the sEH:C9, sEH:A4 and sEH:A8 complexes were determined at 2.17 Å, 2.41 Å and 2.06 Å resolution; the procedure is described in detail in supplement II.

Sample preparation: 380 µL D_2O PBS-buffer with a PKA concentration of 45 µM and additional 150 mM NaCl were combined with 20 µL DMSO-d_6 with appropriate ligand concentration (L1: 300 µM, L2: 900 µM, L3: 1200 µM, L4: 900 µM). 2 mM TCEP was added as an anti-oxidation agent. 380 µL D_2O sodium-phosphate buffer with an sEH concentration of 30 µM were combined with 20 µL DMSO-d_6 with ligand concentrations between 0.25 mM and 1 mM, depending on ligand solubility.

NMR experiments: INPHARMA spectra were recorded using a standard NOESY experiment on a 700 MHz or 900 MHz spectrometer with cryogenically cooled probe heads (Bruker, Karlsruhe). The mixing time was 800 ms for PKA and sEH, and 600 ms for GP. STD spectra were recorded with a standard pulse sequence (stddiffesgp.2) on a 400 MHz spectrometer. Number of scans and points in the time domain are given in the supplement. Saturation was achieved by a train of shaped 90° pulses of 50 ms length. A number of 160 selective pulses was applied, leading to a total length of saturation of 8 s. This duration was used, since the usually applied 2 s did not fully saturate the resonances. The on-resonance irradiation was performed at +0.5 or -2 ppm and off-resonance irradiation was set to +30 ppm.

Molecular docking: For PKA 1000 binding modes per ligand (5 independent runs with 200 modes each time) in the ATP binding site of PKA (rigid protein, crystal structure 3DNE) were created using the docking software PLANTS. These were clustered with 2 Å RMSD similarity filter and the best scoring structure from every cluster was used as a representative structure. During docking the ligand was kept flexible and the protein rigid. It turned out that it is not necessary to create such high numbers of poses for obtaining a high diversity of docking poses. Additionally the number of docking poses per ligands were quite differently distributed. Therefore we decided to use the same number of 10 docking poses per ligand for the systems GP and sEH. In retrospect, it would have been sufficient to use 10 docking poses for PKA as well.For GP 10 docking modes per ligand were created using PLANTS in the crystal structures 1C8L and 1C50. For sEH, 10 docking modes per ligand were created using PLANTS in the crystal structure 1VJ5 and the two structures of opposite binding mode direction with the best scores were chosen.

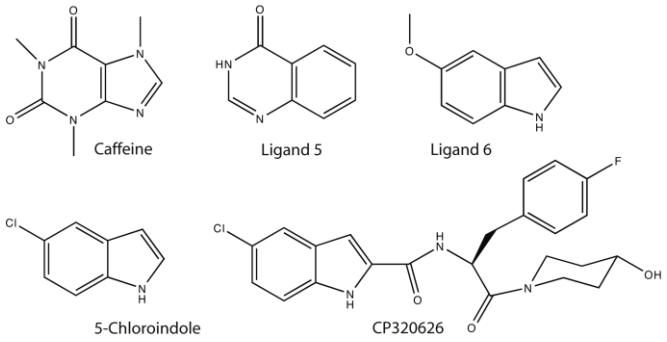
Energy minimization (EM) and molecular dynamics simulation (MD): Calculations were done as implemented in Gromacs[3], using the amber99sb force field[4] for the protein and the General Amber Force Field[5] for the ligand. EM was done with conjugate gradient integrator in 1000 steps. MD was done using stochastic dynamics integrator with a stepsize of 2 fs. Simulations were done for 1 ns in implicit solvent (GBSA) and 500 snapshots were saved for PKA ligands and 100 for GP ligands.

Back-calculation of peak volumes: INPHARMA, trNOE and STD peak volumes were back-calculated with our software SpINPHARMA using the complete-relaxation matrix approach. Methyl-groups were described using a 3-site jump model[6]. STD saturation was applied to all methyl groups. Protons within a distance of 8 Å from any ligand proton were considered. For PKA and sEH ligands the INPHARMA peaks were normalized to the sum of all peaks in the direct dimensions[7]. The correlation times were estimated to be 15 ns for PKA (MW 40 kDa), 100 ns for GP (MW 98 kDa plus dimer formation) and 15 ns for sEH (MW 39 kDa). K_d values were determined by STD competition experiments for PKA (L1: 30 µM, L2: 90 µM, L3: 90 µM, L4: 90 µM) and IC_{50} values for sEH (A4: 3.9 µM, D3: 490 µM, C9: 14.7 µM and A8: 98 µM). For GP all K_d values were set to 100 µM. On-rates (k_{on}) were assumed to be in the diffusion limit 10^8 $M^{-1}s^{-1}$.

Glycogen Phosphorylase

Glycogen phosphorylase (GP) degrades glycogen by phosphorylation of the glucose molecules. Inhibition of GP will therefore lower the blood sugar level, a desired effect in treating type 2 diabetes which makes GP an interesting drug target[8]. GP hosts four different binding sites: the active site, the inhibitor site, the allosteric site and the new allosteric site, making

INPHARMA the ideal tool for determining the binding site of a specific ligand. This was successfully demonstrated by the fact that INPHARMA is observed between binders of the inhibitor site and binders of the new allosteric site respectively, but not between binders of either site[9]. Based on the NOESY and STD spectra of reference[9], we determined the binding modes of the ligands in Scheme 1.



Scheme 1. GP binding ligands: caffeine and ligand 5 bind in the inhibitor site[9,10]. CP320626, 5-chloroindole and ligand 6 bind in the new allosteric site[9,11].

GP exists in many isoforms and for this NMR study the rabbit muscle GPb (inactive) form was used. The structure (PDB code 1C8L[10]) with bound caffeine at the inhibitor site was chosen for docking of ligand 5 to GP. PDB structure 1C50[11] with compound CP320626 bound to the new allosteric site was applied to obtain the protein coordinates for the docking of 5-chloroindole and ligand 6 to the new allosteric site. Interestingly, the ligand CP320626 is an elongated derivative of 5-chloroindole, therefore a similar binding motif can be expected. For docking, the crystal water molecules were kept in the same position as in the crystal structure, as they are stabilized by protein side chains and play an essential role in the binding of caffeine. For ligand 5, the peaks were back-calculated using the crystal structure of bound caffeine (Fig. 2a) as a reference ligand for INPHARMA and result in a chemically plausible binding mode (Fig. 2b) as well as an R_{STI} value similar to the ones found for PKA. The result for 5-chloroindole and ligand 6 (Fig. 2d) is also chemically plausible: The expected binding mode of compound CP320626 (Fig. 2c) is observed and both ligands share a common pharmacophore. The binding modes of two additional GP ligands are also derived and presented in supplement II.

Figure 2. Crystal structure of GP with the inhibitor site (a,b) and new allosteric site (c,d) indicated. Left site: the crystal structures of bound caffeine (a) and CP320626 (c). Right site: the complex structures found by R_{STI} for ligand 5 (b), 5-chloroindole and ligand 6 (d).

References

[1] T. Langer, M. Vogtherr, B. Elshorst, M. Betz, U. Schieborr, K. Saxena, H. Schwalbe, *ChemBioChem* **2004**, 5, 1508-1516;
[2] R. A. Engh, A. Girod, V. Kinzel, R. Huber, D. Bossemeyer, *JBC*, **1996**, 271, 26157-26164;
[3] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theor. Comp.* **2008**, 4, 435-447;
[4] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* **2006**, 65, 712–725;
[5] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comp. Chem.* **2004**, 25, 1157-1174;
[6] J. Tropp, *J. Chem. Phys.* **1980**, 72, 6035-6043;
[7] M. Koeck, C. Griesinger, *Angew. Chem.* **1994**, 106, 338-340; *Angew. Chem. Int. Ed.* **1994**, 33, 332-334;
[8] N. G. Oikonomakos, *Curr. Protein Pept. Sci*. **2002**, 3, 561-586;
[9] a) I. Krimm, *Med. Chem. Comm*. **2012**, 3, 605-610; b) I. Krimm, J.-M. Lancelin, J.-P. Praly, *J. Med. Chem*. **2012**, 55, 1287-1295;
[10] K. E. Tsitsanou, V. T. Skamnaki, N. G. Oikonomakos, *Arch. Biochem. Biophys.* **2000**, 384, 245-254;
[11] N. G. Oikonomakos, V. T. Skamnaki, K. E. Tsitsanou, N. G. Gavalas, L. N. Johnson, *Structure* **2000**, 8, 575-584;

Supplement II

Part I: Protein kinase A (PKA)

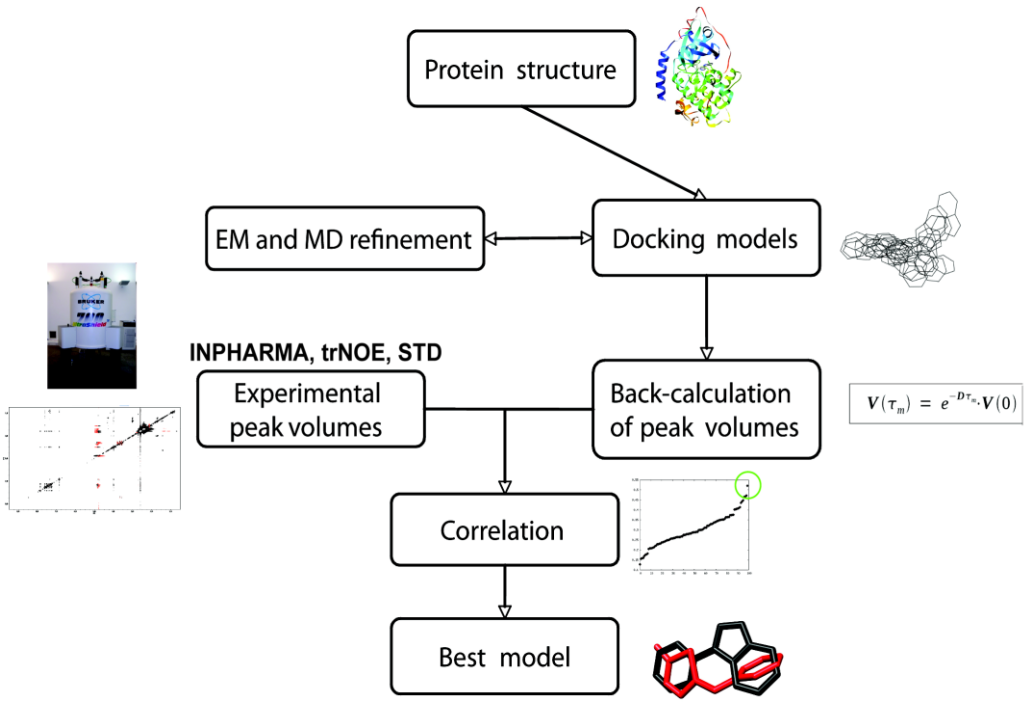


Figure S1: The flowchart gives an overview of the NMR-based binding mode determination methodology. Docking models are created starting with a crystal structure or homology model. These can be further optimized with energy minimization (EM) or molecular dynamics simulations (MD). STD, trNOE and INPHARMA peak volumes for all pairs of docking poses are back-calculated with the software SpINPHARMA. Then the back-calculated volumes are correlated with the experimental NMR data and the best scoring model is chosen. The scores of STD, trNOE and INPHARMA are combined to STI as seen in the equation in the main text.

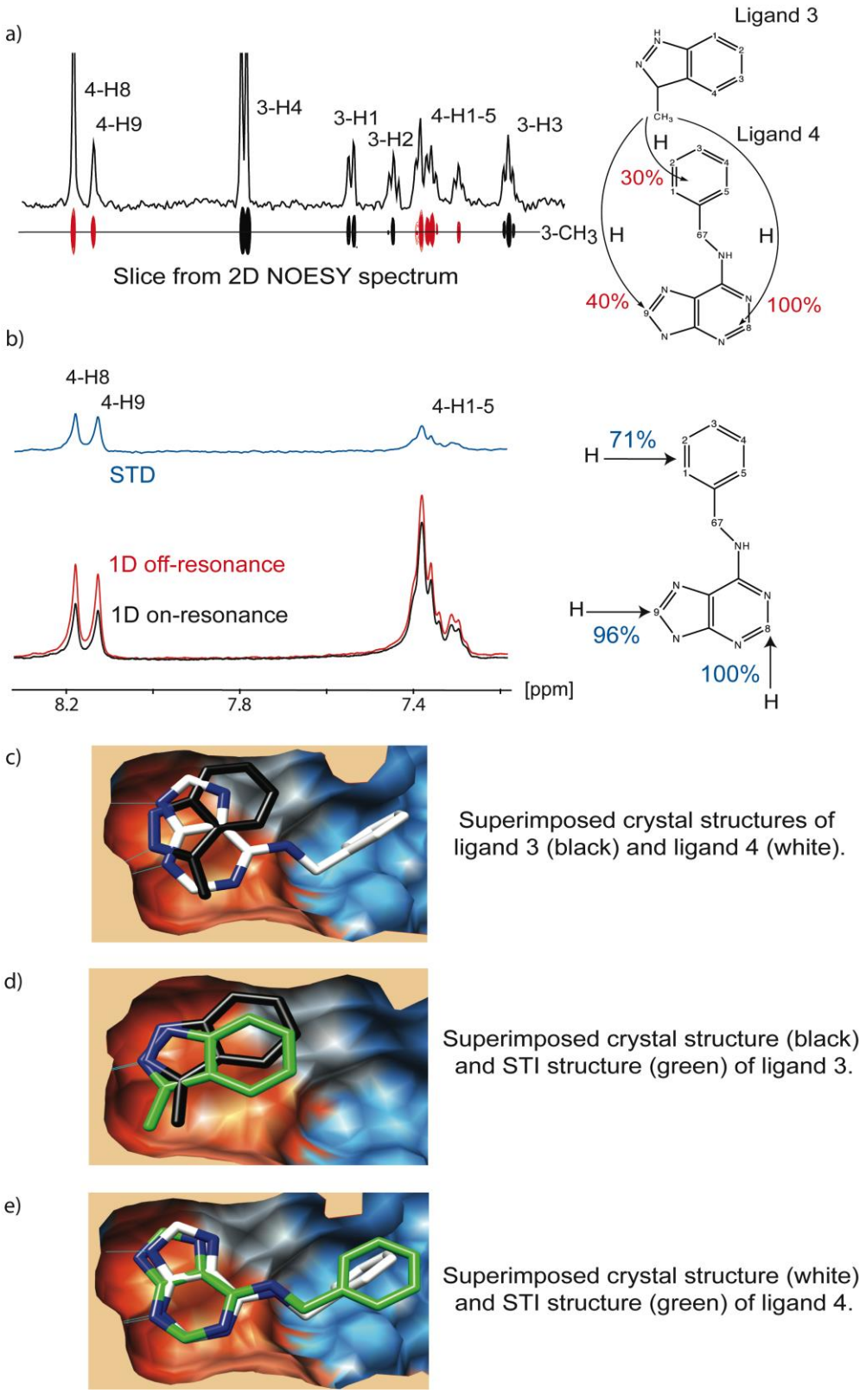


Figure S2: INPHARMA spectrum for the pairwise combination of ligands 3 and 4 as well as STD spectrum of Ligand 4 are shown. a) The slice through the NOESY spectrum at the resonance of the methyl group of ligand 3 in ω_1 demonstrates that the magnetization is transferred via protein protons (H) to ligand 4 and indicates e. g. a closer spatial proximity of 3-CH_3 to 4-H8 than to 4-H1-5. b) With STD, the protein protons are saturated and this deviation of magnetization from Boltzmann

equilibrium diffuses to the bound ligand via NOE. Ligand protons in close contact to the protein such as 4-H8 and 4-H9 are more efficiently saturated via protein protons than solvent exposed protons such as 4-H1-5. The STD sample contained 45 µM PKA and 900 µM of ligand 4. The INPHARMA sample contained additionally 1200 µM of ligand 3. c) The crystal structures of ligand 3 (black) and ligand 4 (white) are superimposed to put the correct prediction of INPHARMA, as well as STD, into evidence. The aromatic ring of ligand 4 is oriented towards the solvent and protons 4-H8 and 4-H9 are found to be buried deep in the binding pocket. d) shows the superimposition of the crystal structure (black) of ligand 3 and the structure determined with our STI methodology (green) and e) shows the superimposition of the crystal structure (white) of ligand 4 and the structure determined with our STI methodology (green).

The scoring of INPHARMA, trNOE and STD data

The Pearson correlation coefficient *R* is used to score the fitting between the experimental and back-calculated INPHARMA, trNOE and STD data. *R* is defined as:

$$R = \frac{\sum_i^{peaks}(V_{exp,i} - \bar{V}_{exp})(V_{calc,i} - \bar{V}_{calc})}{\sqrt{\sum_i^{peaks}(V_{exp,i} - \bar{V}_{exp})^2 \sum_i^{peaks}(V_{calc,i} - \bar{V}_{calc})^2}}$$

R is not squared in order to differentiate a positive correlation from a negative correlation. This is especially important in the case of STD, because here a correct binding mode has a positive correlation, while a wrong binding mode which is flipped by 180° would have a negative correlation.

We also applied a quality factor Q, defined as:

$$Q^2 = \frac{\sum_i^{peaks}(V_{exp,i} - V_{calc,i})^2}{\sum_i^{peaks}(V_{exp,i})^2}$$

In the case of INPHARMA we found that Q is very sensitive to small changes in the geometry, while *R* is less sensitive. Therefore we use *R* in a first stage, where we want to find the correct binding mode out of docking structures. In a second step a geometry optimization can be done using energy minimization and molecular dynamic simulations (*vide infra*). These structures are then scored by *R*

and Q. In the figures S10 and S11 the optimized structures of the ligand pairs 1&2 and 3&4 are shown.

The combination of INPHARMA, trNOE and STD data

R_{STD}, R_{trNOE} and $R_{INPHARMA}$ are combined to R_{STI} as described in the main text by averaging. The weighting of the different parameters is chosen depending on the error of integration of the experimental peak volumes, i. e. the signal to noise ratio. For the INPHARMA peaks we integrated the peaks with random areas of the noise and estimated the error to be between 10 % and 20 %. For the intraligand (trNOE) peaks we estimated 5% and for the diagonal peak volumes 2%. For the STD spectra we did the same and come to a similar error estimation between 10 % and 20 %. The intraligand peaks have the lowest error, but also contain the least information, especially with rigid small molecules. Therefore it is justified to apply a 1:1:1 weighting. If the spectra of either the INPHARMA NOESY or STD have a low S/N ratio (e.g. 5), it is advisable to use another weighting scheme.

In the main text it was mentioned that STD shows binding for a wider range of unbinding rates than INPHARMA. Indeed, STD requires only one unbinding-event during the saturation lasting several seconds, while INPHARMA requires at least two off-events in the mixing time of a few 100 ms. When STD is recorded, the K_d values of the ligands can be estimated. Given that the K_d value of a ligand is lower than 10 µM and still shows INPHARMA signals, we recommend to be careful that INPHARMA and STD really emerge from the same binding site. One can imagine that the STD signals show a strong ligand binding event and the INPHARMA signals a weaker binding event to another binding site. To make sure that an effect of a strong binding site for STD does not occur, the K_d values of every ligand should be determined. This can be done with STD if a reference ligand exists. If no reference ligand is available, other biophysical methods such as isothermal titration calorimetry (ITC) or surface plasmon resonance (SPR) could be used. Given that there are ligands with a binding affinity below 10 µM, it can be assumed that they are strong binders and potentially invisible to INPHARMA. One option to handle ligands with binding affinities below 10 µM is to titrate them with weaker binding ligands (typically in the 100-1000 µM range) during an STD experiment. When the signal intensity of the ligand decreases, it binds the same binding pocket as the weaker binder. Then the INPHARMA signal most likely reflects competitive binding in the same pocket.

Multiplexing of several ligands

In order to evaluate the impact of having more ligands than two (multiplexing), the R_{STI} scores of all ligand pairs were combined to $R_{STI}^{Multiplexing}$. As seen in Fig. 2 in the main text, the differentiation between the correct structures and those with larger RMSD is even more pronounced and the correct binding mode can be identified with Multiplexing more easily. In the same way as STD, trNOE and INPHARMA together contain more information as either of them alone, the combination of several ligand pair contains more information as one single ligand pair. Additionally experimental errors are reduced as different data sets are used for the different ligand pairs. Interactions that are seen for one pair are probably not seen for another. The fact that the ligands are different in size and

volume leads to even more information which part of which ligand overlaps in a superimposition. Obviously Multiplexing can only improve INPHARMA and not STD and trNOE, as only INPHARMA uses pairwise ligand combinations.

Refinement by molecular dynamics simulations

In the main text we describe the methodology to derive ligand binding modes by a combination of easy to implement NMR-based methods. First we recommend to record INPHARMA, trNOE and STD spectra to choose the best scoring binding mode with R_{STI}. If more than two ligands are available, the reliability of complex structure determination can be improved even further by applying $R_{STI}^{Multiplexing}$. MD refinement of the determined ligand binding modes is recommended as described in the following.

In order to cover a larger conformational space, docking structures with good R_{STI} values can be refined by MD simulation. By taking snapshots of the resulting trajectories many new structures close to the starting structure are generated. These can then be cross-validated against the experimental data in order to find the best scoring one (see figures S10-11).

In our investigations, a MD simulation for 1 ns was sufficient to cover the conformational space and to obtain structures with excellent correlations with the experimental data. In the case of PKA, the correct structures, i.e. structures close to the crystal structures of the respective complexes, remained stable in the binding pocket. On the other hand, wrong structures, i.e. structures which had an RMSD to the crystal structures of more than 5 Å left the binding site more easily. In general we find that for wrong structures the R factors drop strongly during the MD simulation or the ligand leaves the binding site. This is an indication that such structures should not be used as complex structures.

In the case of sEH the binding site has the shape of a tunnel, and the ligands can leave easily to both sites. This is indeed observed during the MD simulation. The most stable structures, which are the ones that stay at the binding site for the longest time, were the ones chosen by R_{STI} or $R_{STI}^{Multiplexing}$ which supports the correctness of these complex structures.

STD peak volumes

To obtain the STD peak volumes V_{STD}, the spectra with off-resonance saturation and on-resonance saturation where integrated and combined according to the equation:

$$V_{STD} = \frac{V_{off\text{-}resonance} - V_{on\text{-}resonance}}{V_{off\text{-}resonance}}$$

Back-calculated peak volumes were treated in the same way.

For the graphical representation of STD in percent, as shown in this supplement, the largest STD peak volume $V_{Max.STD}$ was set to 100% and the other peak volumes were normalized according to:

$$STD_{in\%} = \frac{V_{STD}}{V_{Max.STD}} * 100$$

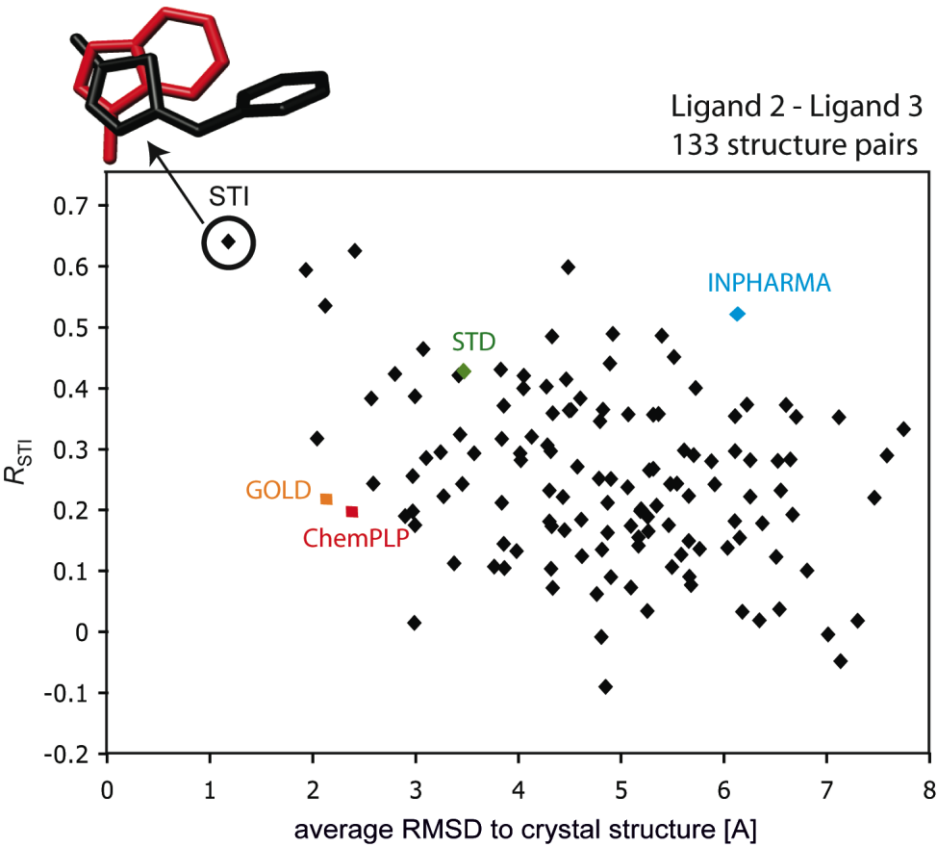


Figure S3: R_{STI} plotted against the average ligand RMSD of ligand 2 and 3 in PKA structures compared to the crystal structures. The best scoring structures selected by the various methods are represented in blue (INPHARMA), green (STD), red (ChemPLP) and yellow (GOLDScore). The combination R_{STI} finds the correct docking mode (circled in black), i. e. the mode with the smallest RMSD to the average crystal structure.

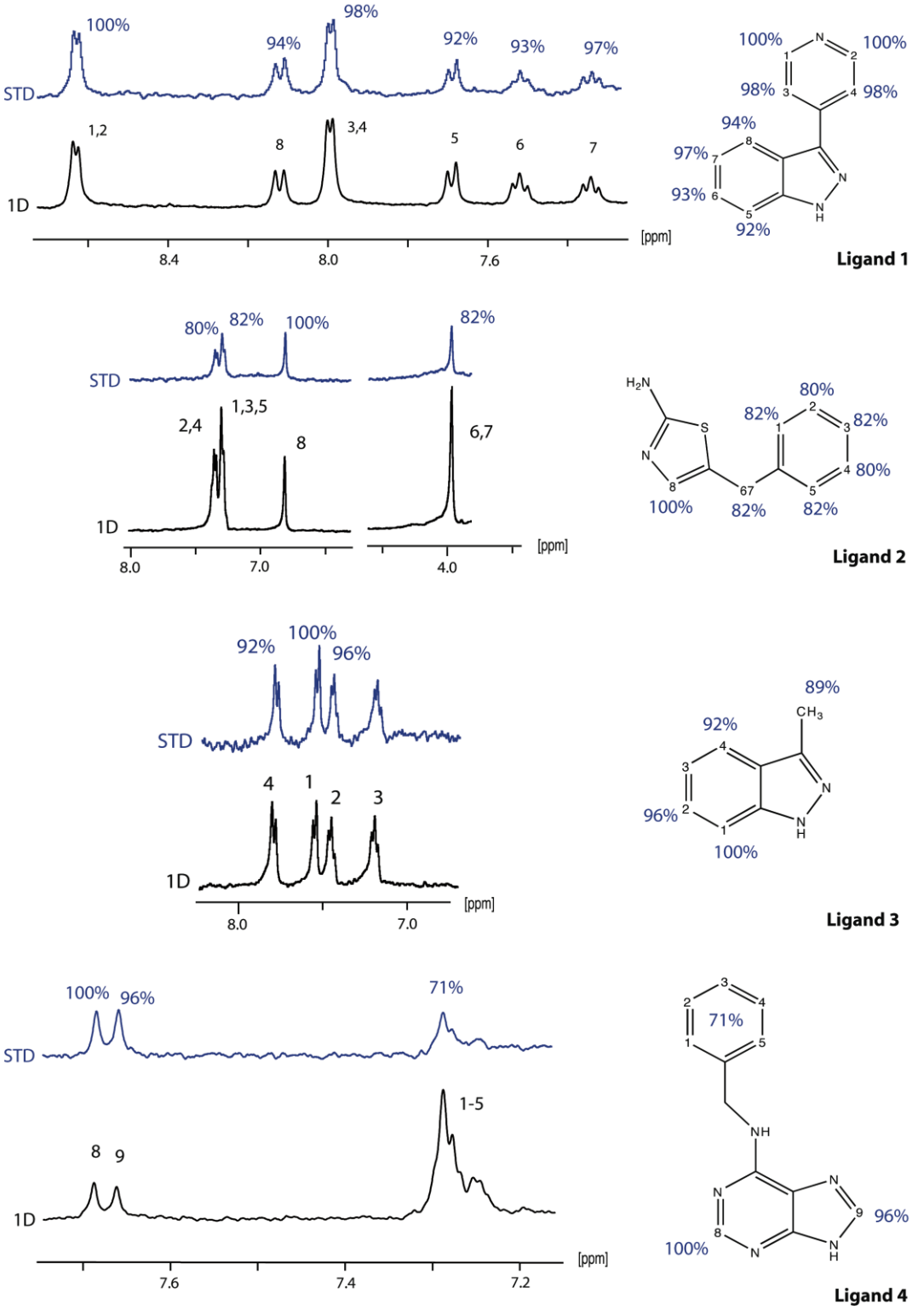


Figure S4: Experimental STD data of the four PKA binding ligands. The normal 1D proton NMR spectrum in the presence of 45 µM PKA is shown in black and the STD spectrum in blue. Ligand concentrations were 300 µM (ligand 1), 900 µM (ligand 2), 900 µM (ligand 3) and 900 µM (ligand 4). Saturation was applied for 8 s on 0.5 ppm on a 400 MHz spectrometer (number of scans = 8; temperature = 298 K; time domain = 16384 points). Proton 3 of ligand 3 was dismissed due to broadening, caused by an overlap with broadened protein signals.

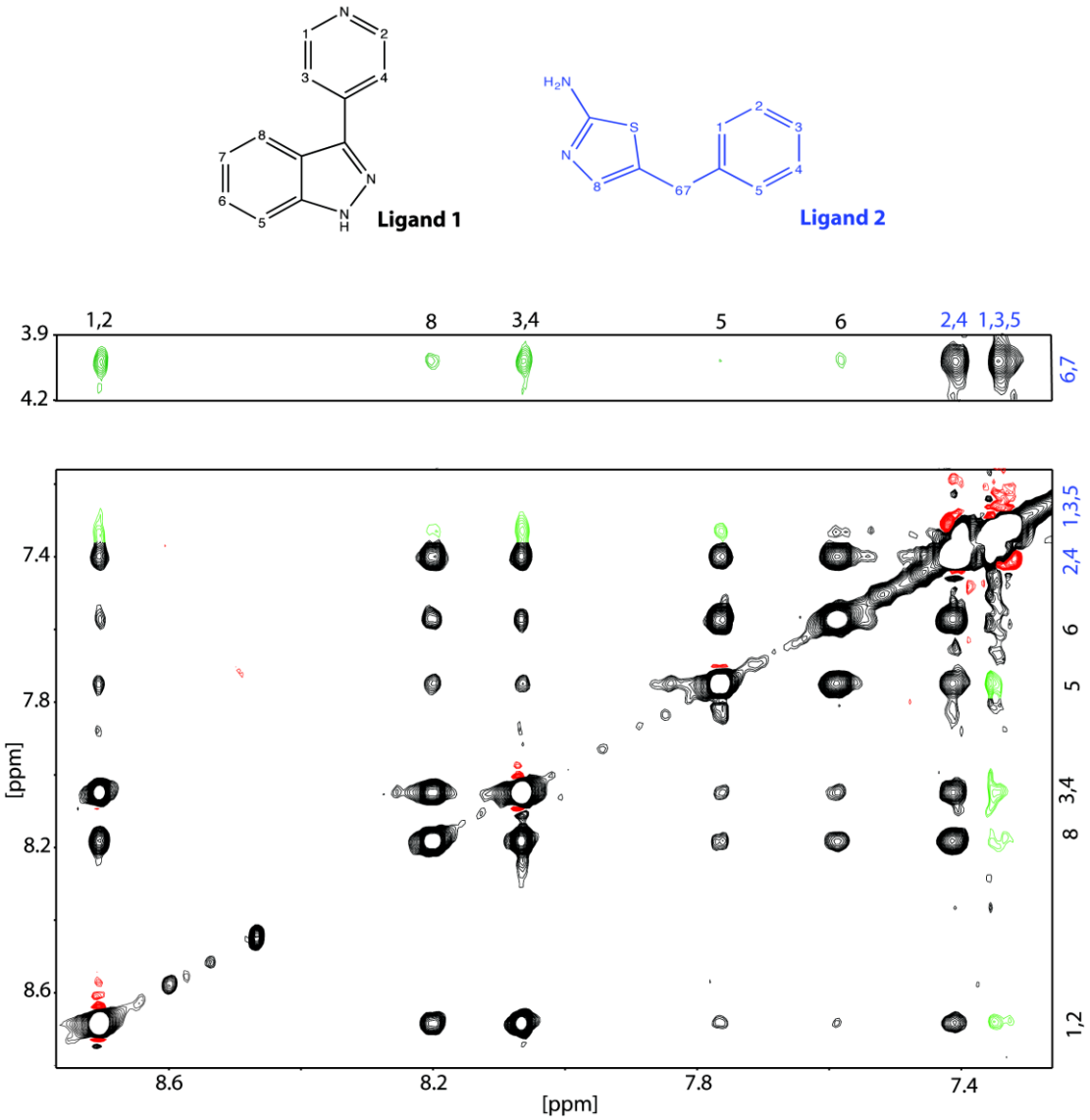
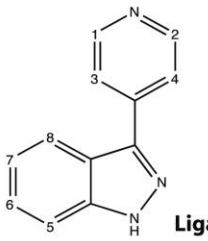
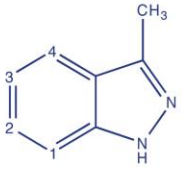


Figure S6: NOESY spectrum of ligand 1 (300 µM) and ligand 2 (900 µM) in the presence of PKA (45 µM). INPHARMA peaks chosen are marked in green. The mixing time was 800 ms on a 700 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 2048 points in F2 and 580 points in F1.



Ligand 1



Ligand 3

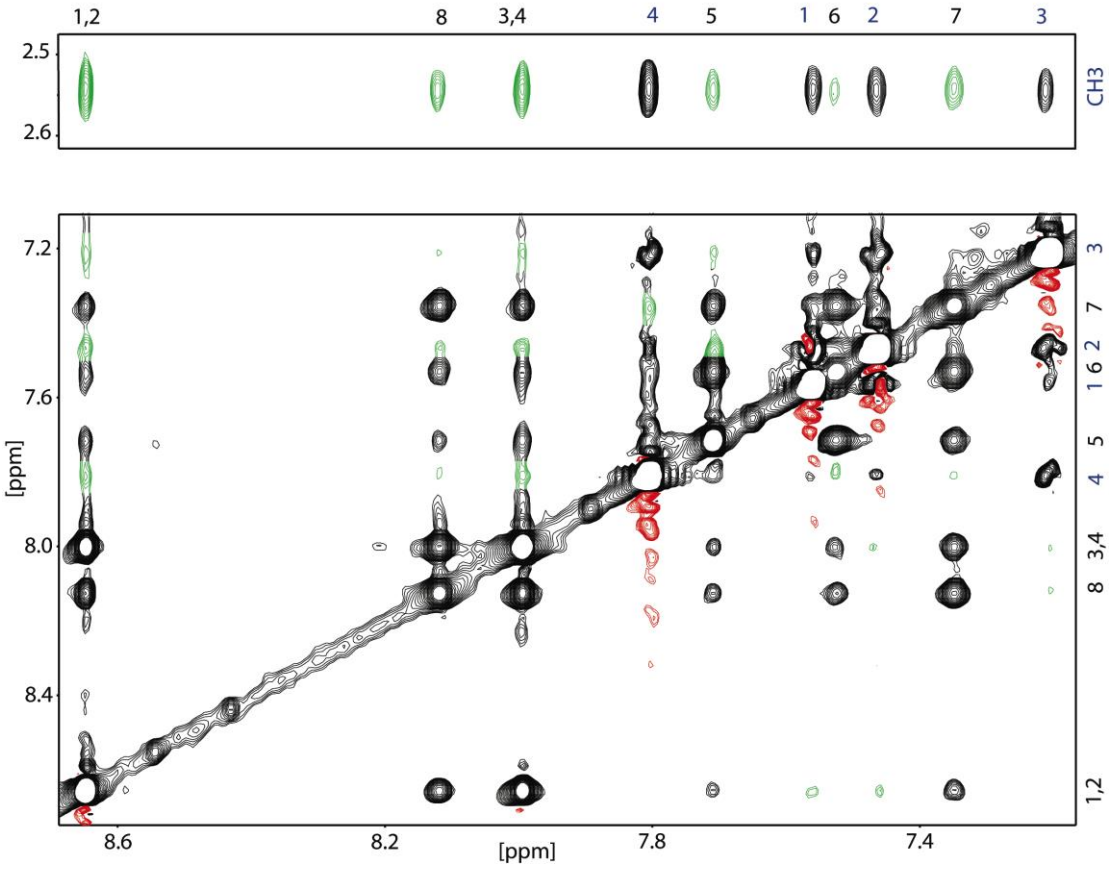


Figure S7: NOESY spectrum of ligand 1 (300 µM) and ligand 3 (900 µM) in the presence of PKA (45 µM). INPHARMA peaks chosen are marked in green. The mixing time was 800 ms on a 900 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 2048 points in F2 and 580 points in F1.

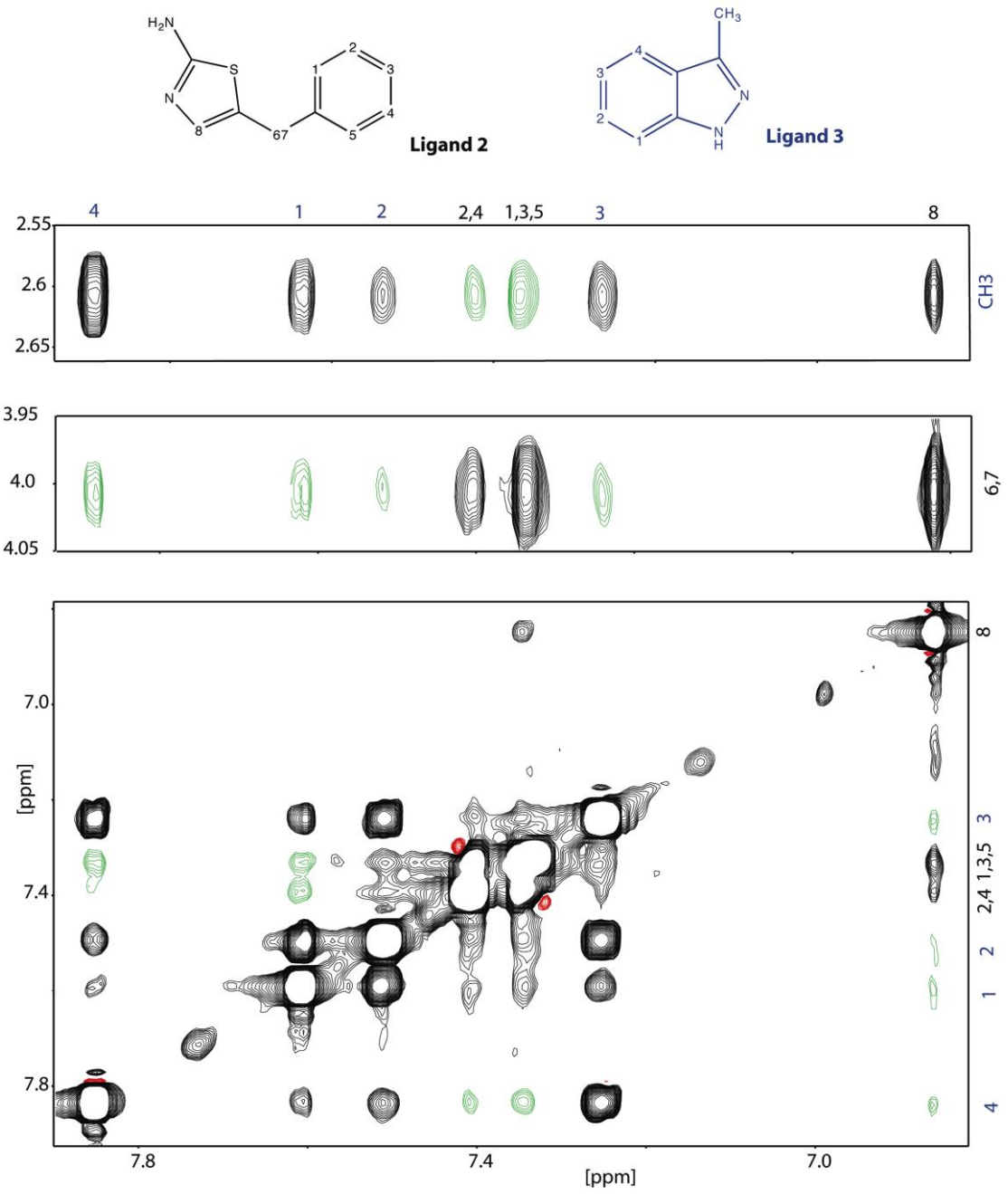


Figure S8: NOESY spectrum of ligand 2 (900 µM) and ligand 3 (1200 µM) in the presence of PKA (45 µM). INPHARMA peaks chosen are marked in green. The mixing time was 800 ms on a 700 MHz

spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 2048 points in F2 and 580 points in F1.

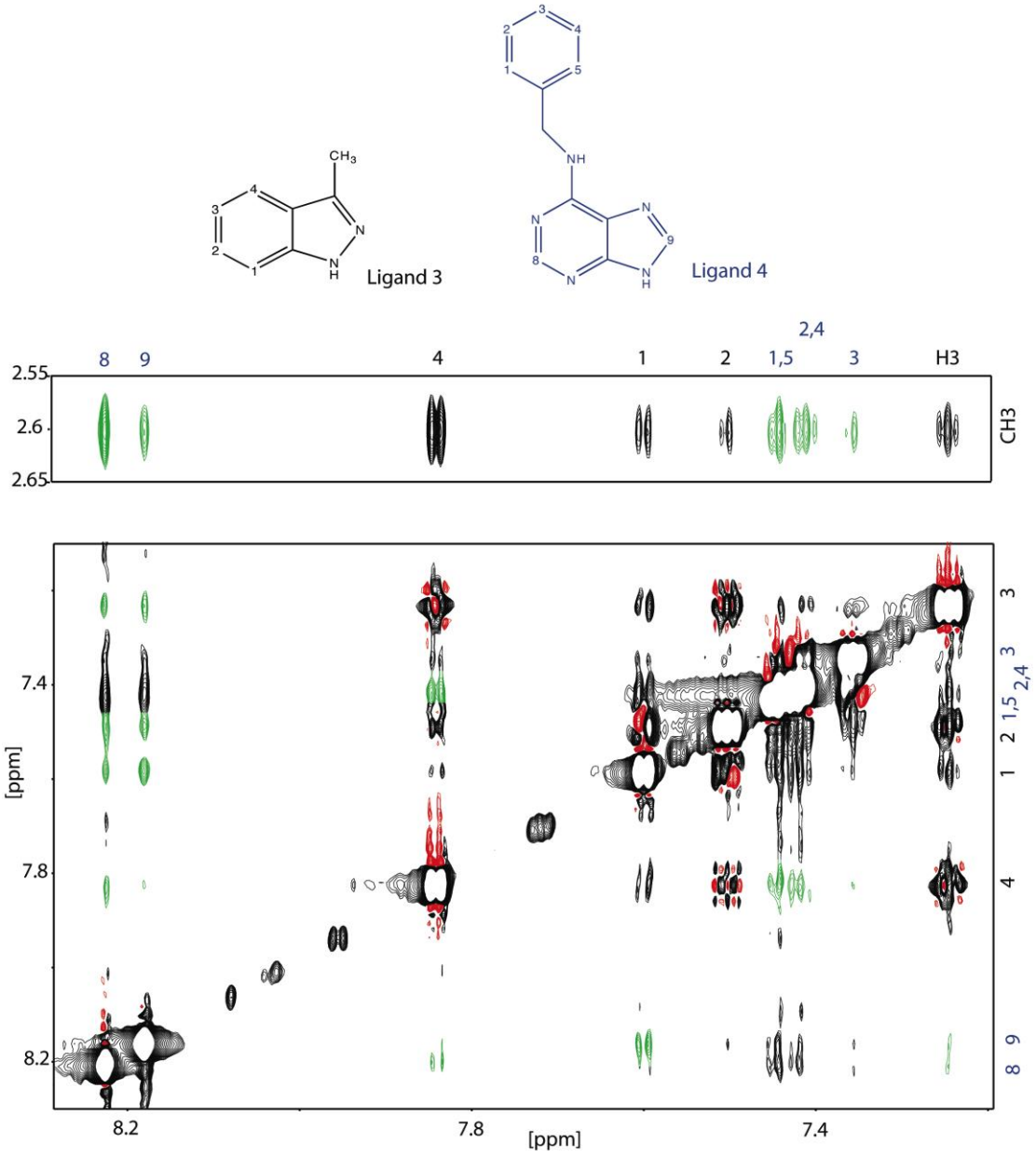


Figure S9a: NOESY spectrum of ligand 3 (900 µM) and ligand 4 (900 µM) in the presence of PKA (45 µM). INPHARMA peaks chosen are marked in green. The mixing time was 800 ms on a 700 MHz

spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 4096 points in F2 and 580 points in F1.

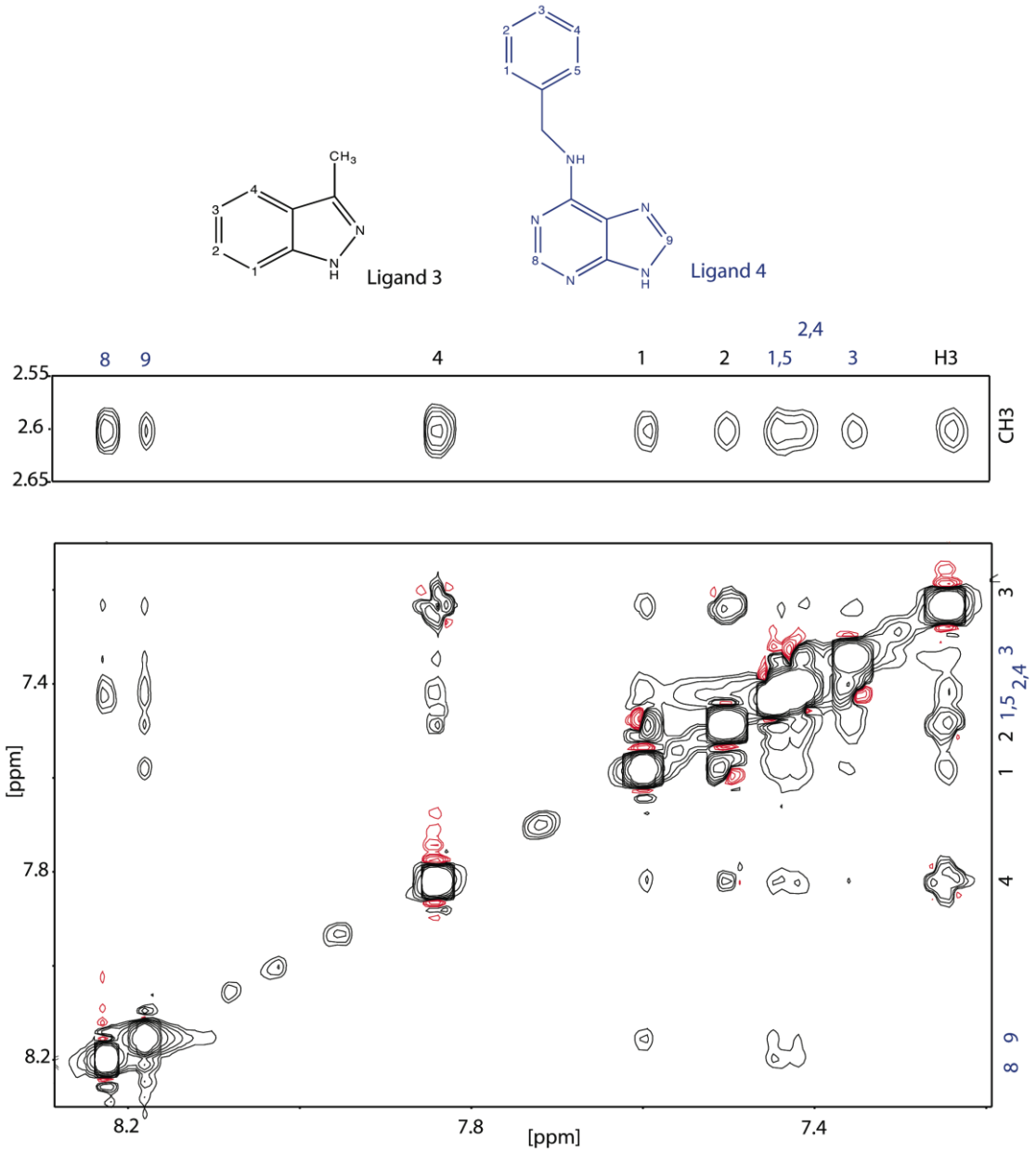


Figure S9b: NOESY spectrum of ligand 3 (900 µM) and ligand 4 (900 µM) in the presence of PKA (45 µM). The processing was done with only 2048 points to demonstrate that the digitization is the only difference to the other spectra. The mixing time was 800 ms on a 700 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 4096 points in F2 and 580 points in F1.

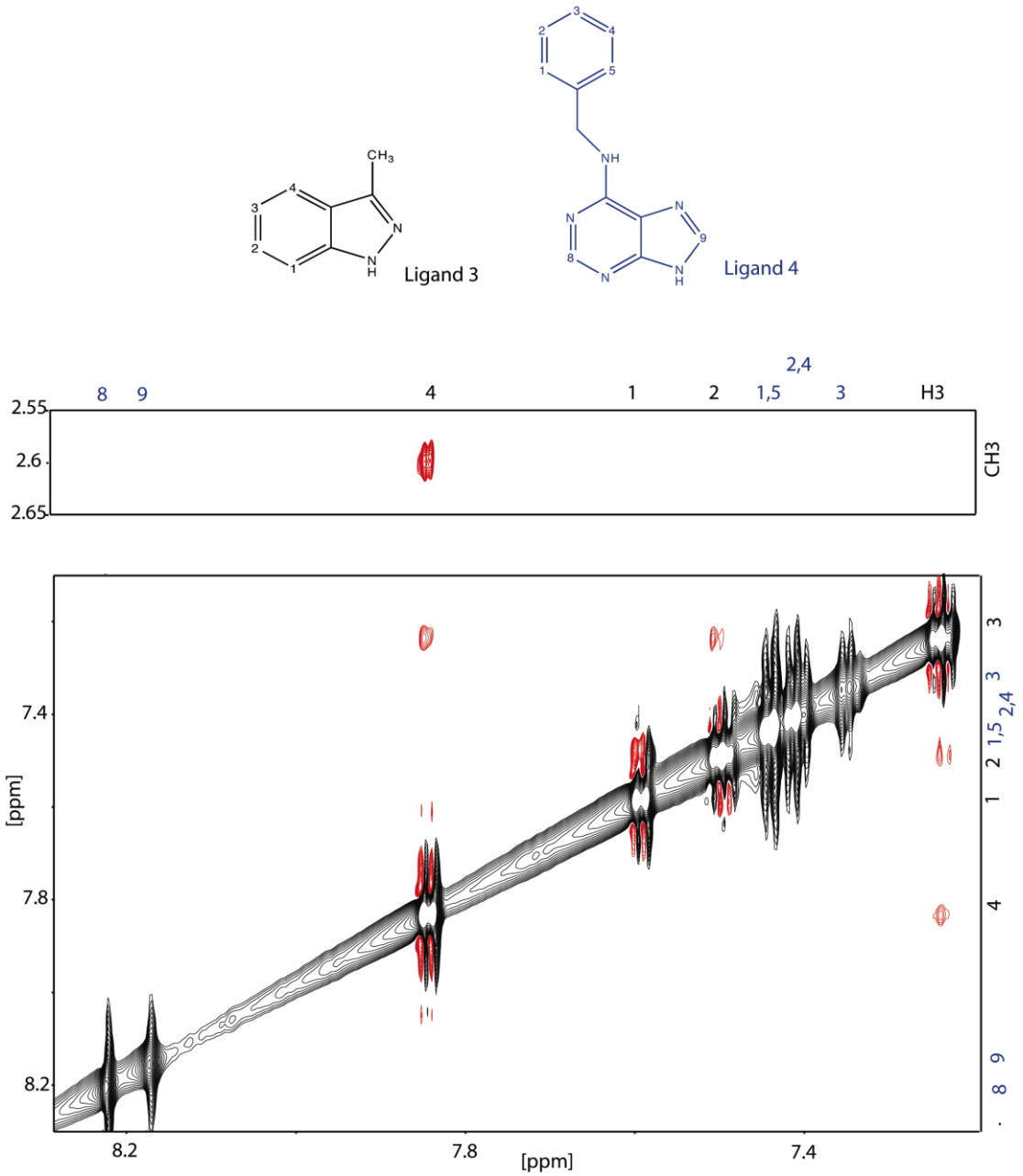
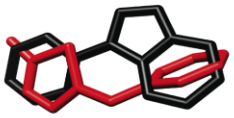
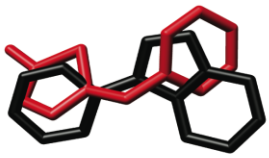


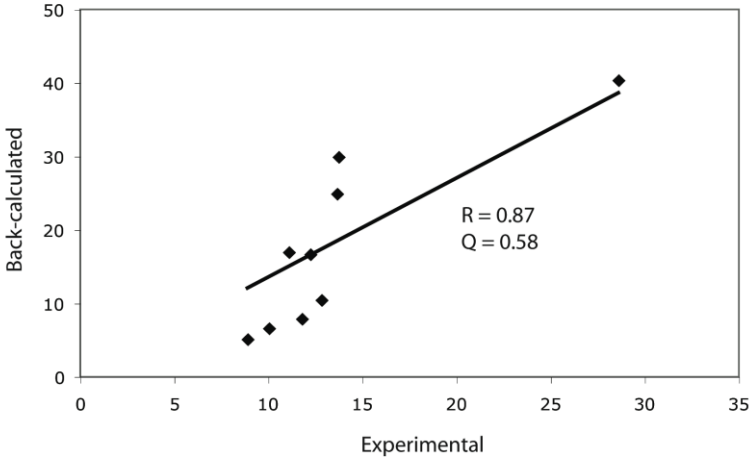
Figure S9c: NOESY control spectrum of ligand 3 (900 µM) and ligand 4 (900 µM) in buffer. INPHARMA peaks and trNOE peaks can not be observed. The mixing time was 800 ms on a 700 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 4096 points in F2 and 384 points in F1.

Ligand pair 1 and 2

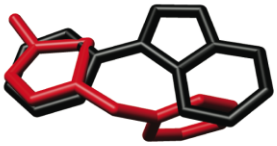


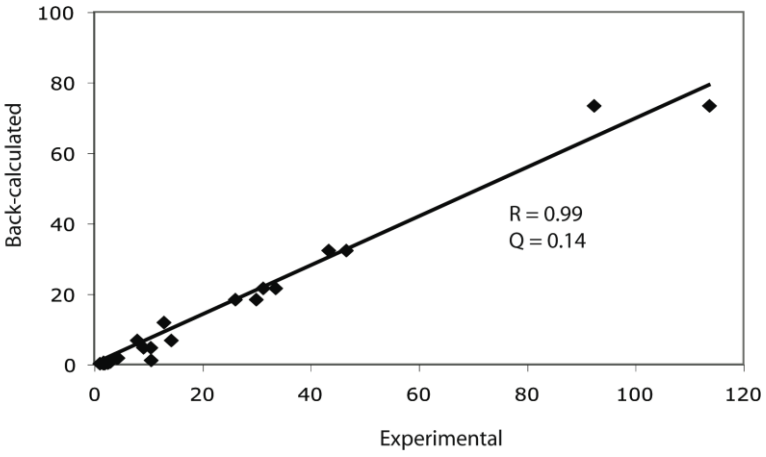
STD peak volumes



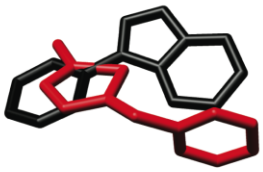


Tr-NOE peak volumes





INPHARMA peak volumes



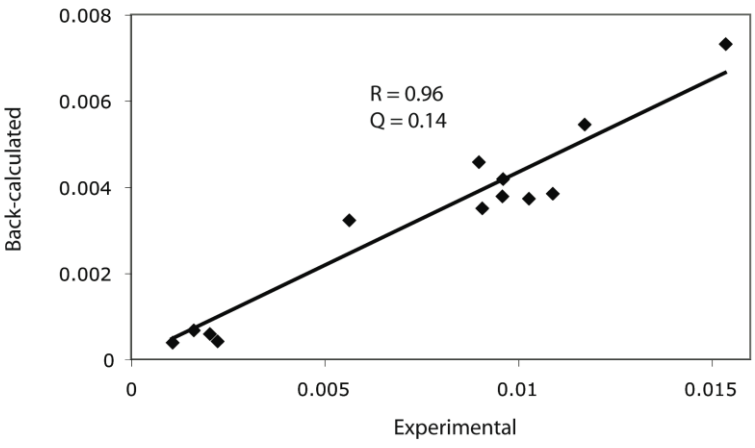


Figure S10: Correlation of the experimental and back-calculated peak volumes of the ligand pair 1 and 2 with the best R_{STI} score, refined during an MD simulation against STD, trNOE and INPHARMA peak volumes, respectively.

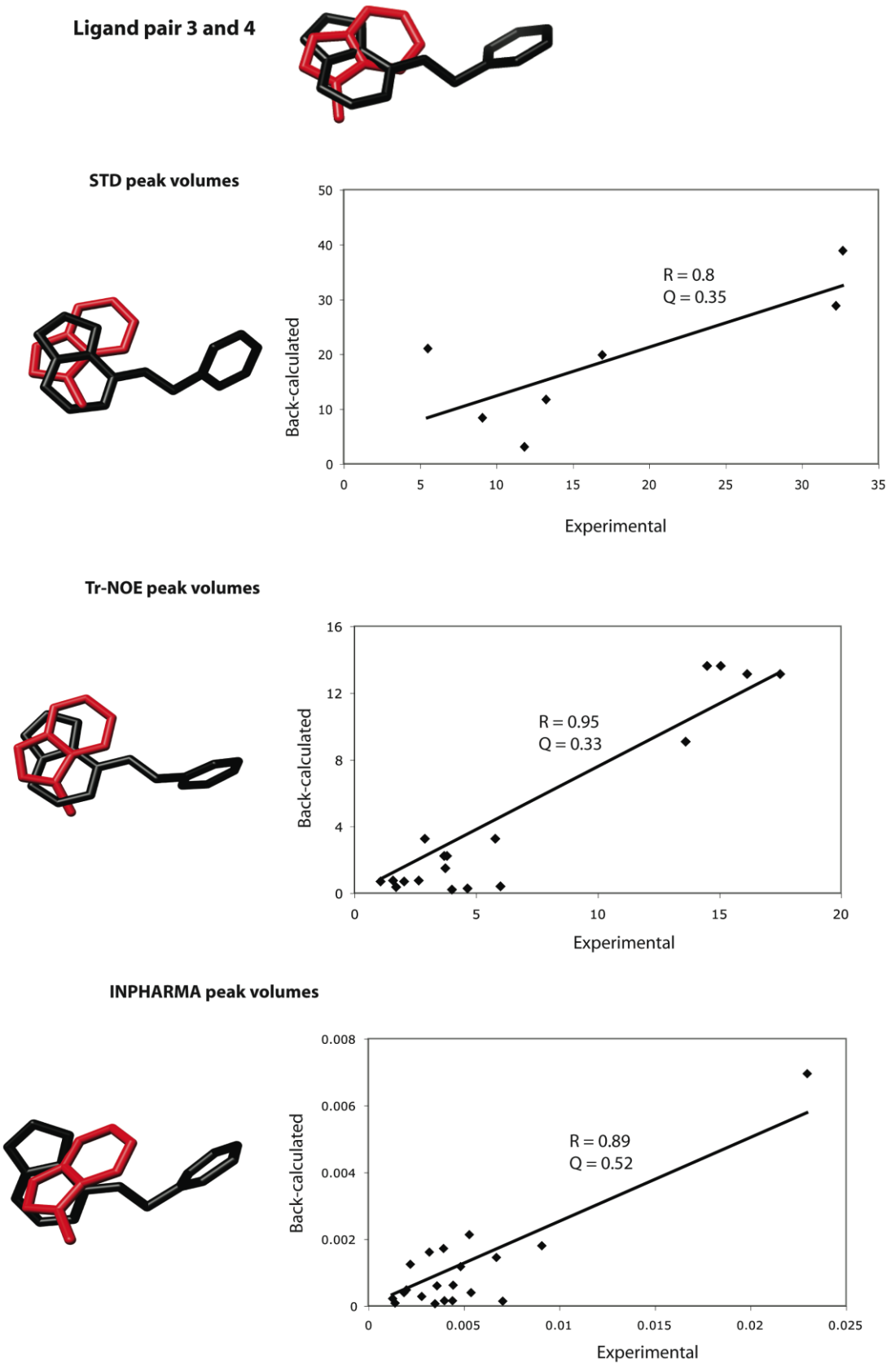


Figure S11: Correlation of the experimental and back-calculated peak volumes of the ligand pair 3 and 4 with the best R_{STI} score, refined during an MD simulation against STD, trNOE and INPHARMA peak volumes, respectively.

Protein side chain flexibility and backpocket penetration - Ligand 9

The ATP binding site of PKA hosts a backpocket, which is guarded by the gatekeeper residue Met[120]. Most kinases display a similar motive, that a backpocket is blocked by a gatekeeper residue. To obtain the correct binding mode of a backpocket penetrating ligand, the gatekeeper residue has to change its conformation. In the following, it will be shown that the STI methodology can handle this issue. Therefore ligand 9 was chosen, from which it is known from the crystal structure, that a penetration into the backpocket takes place.

First, a STD spectrum was recorded (Fig. S12). Visual inspection of the STD efficiencies suggests that protons H1-H4 are in contact with protein protons, while the protons H9-H12 are exposed to the solvent. The ligand resembles partially ligand 1, yet the binding mode seems to be different, as the protons H6 and H7 exhibit much stronger STD effects in comparison to protons H5 and H8. In ligand 1, protons H5 to H8 had similar STD efficiencies.

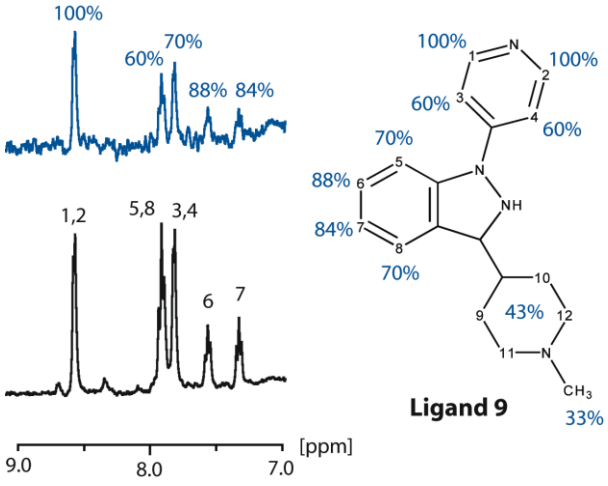


Figure S12: Experimental STD data of PKA binding ligand 9. The normal 1D in the presence of 30 µM PKA is shown in black, the difference spectrum between 1D and STD in blue. Ligand concentration was 0.5 mM. Saturation was applied for 8 s on 0.5 ppm on a 400 MHz spectrometer (number of scans = 8; temperature = 298 K; time domain = 16384 points).

Second, an INPHARMA NOESY was recorded. Ligand 3 was chosen as the second ligand, as there were fewest overlapping peaks in the 1D spectrum with this ligand compared to ligands 1,2 and 4. Visual inspection of the spectrum confirms the STD observations: The protein buried methyl group of ligand 3 is very close to protons H1-H4, but far from the protons H9-H12 of ligand 9. 10 docking modes of ligand 9 were created and STD, trNOE and INPHARMA were back-calculated for all docking

modes. The result can be seen in Fig. S13: Two docking modes, very close in RMSD, score better than the other structure pairs. If it is assumed that the correct binding mode of ligand 3 is already known (see above), the selection gets even clearer. If the obtained binding mode is inspected and compared with the crystal structure (Fig. S14) it can be seen that they are very close in RMSD. Yet, the backpocket can not be entered, as the gatekeeper residue Met^{120} is in the closed and not in the open conformation. Therefore it would not be noticed, that this ligand penetrates into the backpocket and has the potential to be optimized to a highly specific binder.

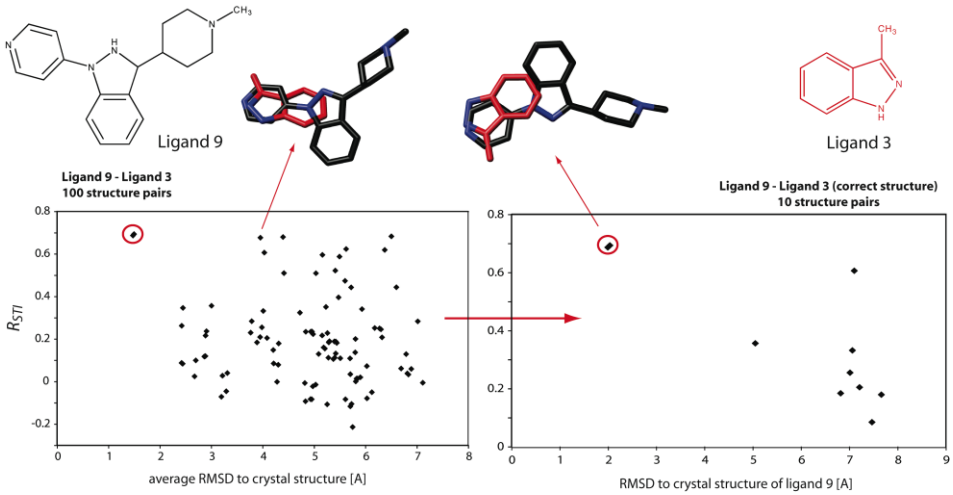


Figure S13: The 100 pairwise combinations of ligand 9 (black) and ligand 3 (red) scored against R_{STI} on the left side. On the right side only the pairs with the correct structure of ligand 3 are shown. Therefore only the RMSD of ligand 9 to the crystal structure is shown and not the RMSD averaged over the coordinates of ligand 3 and ligand 9.

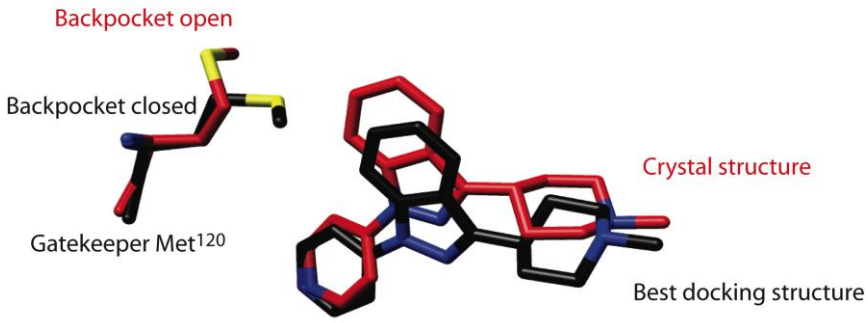


Figure S14: Best docking model of ligand 9 by R_{STI} in black and the crystal structure in red. The gatekeeper residue Met^{120} normally closes the backpocket in PKA. But for ligand 9 in the bound state, the backpocket is open as seen in the crystal structure.

Then, the obtained complex structure was run in an unrestrained MD simulation for 1 ns, as was done in the MD refinement described above. During the MD simulation the gatekeeper immediately opens and ligand 9 penetrates into the backpocket. And most interestingly, these structures with the backpocket penetration score (best $R_{INPHARMA}$: 0.94) better than the initial docking model with the closed backpocket (best $R_{INPHARMA}$: 0.86). The structure with the best score is shown in Fig. S15. The RMSD of this complex structure to the crystal structure of 1.3 Å is reduced compared to the complex structure with the gatekeeper in the closed conformation (RMSD: 1.9 Å). Upon MD we make the interesting observation that the ligand can enter the backpocket, thus reproducing the crystal structure, and that this conformation scores higher than the previous one without MD.

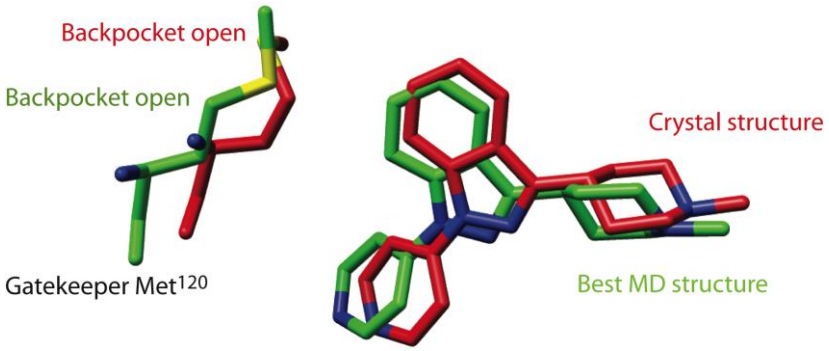


Figure S15: Best MD structure of ligand 9 by R_{STI} in green and the crystal structure in red. The gatekeeper residue Met^{120} is in the open position in both structures.

Protein flexibility and ligand torsion angles - Ligand 10

The second example presents ligand 10 and treats two problems that occur within the docking process of this ligand, when compared with the crystal structure of the complex of PKA and ligand 10: i) the ligand has two rotatable bonds which are not correctly rotated by the docking software and ii) a side chain Asp^{184} blocks the correct orientation of the solvent exposed part of the ligand.

First, a STD spectrum was recorded (Fig. S16), and the binding epitope of the ligand could be clearly identified. Protons H1 and H2 are buried in the protein, while the C6-bound methyl group is exposed to the solvent. Strong differences in STD efficiencies between protons H1 and H3, as well as H4 and H5, indicate that the ligand protons are differentially exposed to the protein surface.

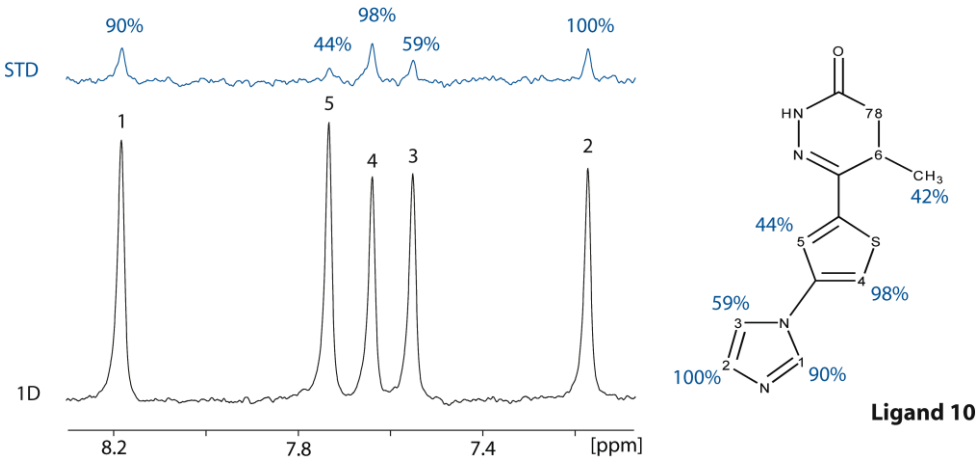


Figure S16: Experimental STD data of PKA binding ligand 10. The normal 1D in the presence of 45 µM PKA is shown in black and the STD spectrum in blue. Ligand concentration was 1 mM. Saturation was applied for 8 s on 0.5 ppm on a 400 MHz spectrometer (number of scans = 8; temperature = 298 K; time domain = 16384 points).

An INPHARMA NOESY was recorded with ligand 2 as the competing ligand. The INPHARMA peaks were back-calculated for all 10 docking poses. The structure of ligand 10 with the best R_{STI} score of only 0.29 is shown in Fig. S18a. It can be seen that the orientation of the ligand is correct. Yet, the ligand structure has an RMSD of 3.49 Å to the crystal structure. The reason for this is, that the methyl-group bearing aromatic ring of ligand 10 cannot rotate into its correct position, because it is hindered by the conformation of the side chain Asp^{184}. It can be seen in the crystal structure, that such a conformational change has to take place to allow the ligand to adopt the correct binding mode.

Therefore we took the same approach as with ligand 9 and determined the bound conformation of ligand 10 by transfer NOESY (Fig. S17). The spectrum clearly indicates that the protons i) H1 and H4, ii) H3 and H5, iii) H5 and the C6-attached methyl group, are close in space in the bound conformation. A NOESY spectrum of ligand 10 in solution without PKA (Fig. S17c) shows no NOEs peaks between the protons H1-H3 and H4-H5, which means that the bond between them can freely rotate. CH3 and H6 show a strong NOE to H5, which means that they are close in space.

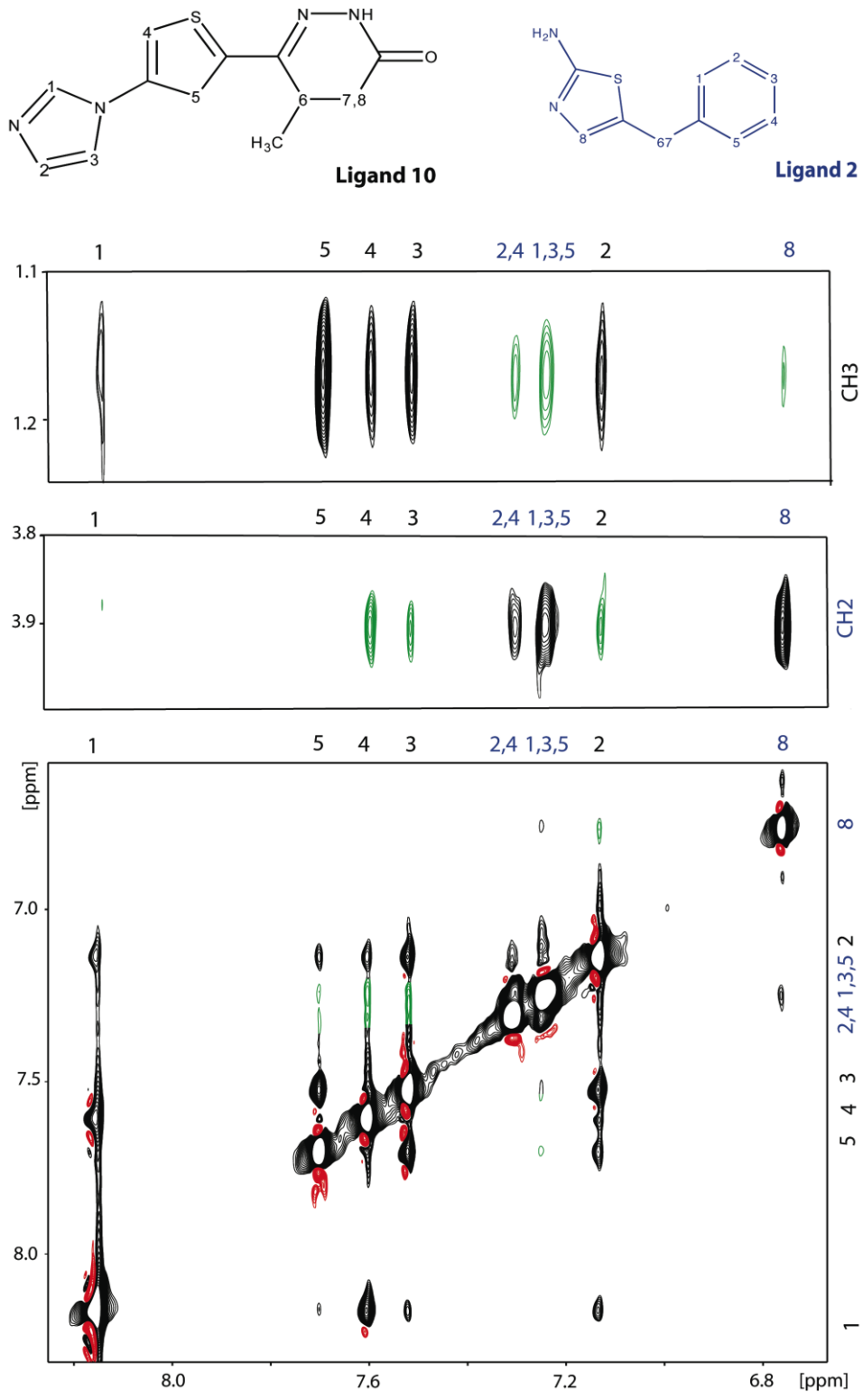


Figure S17a: NOESY spectrum of ligand 10 (1 mM) and ligand 2 (1mM) in the presence of 45 µM PKA. INPHARMA peaks chosen are marked in green The mixing time was 800 ms on a 700 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 96 scans, 4096 points in F2 and 288 points in F1.

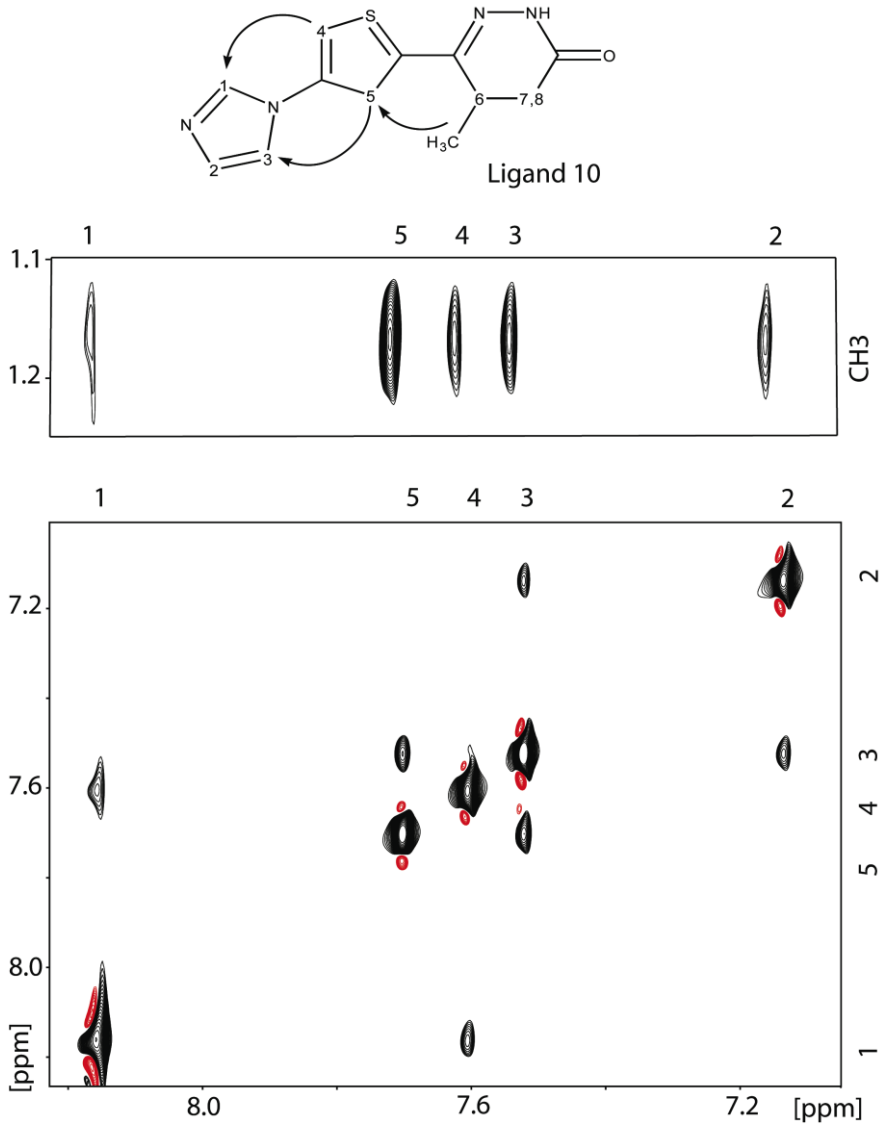


Figure S17b: NOESY spectrum of ligand 10 (1 mM) in the presence of 45 µM PKA. The bound conformation of ligand 10 can be derived from trNOE. The mixing time was 800 ms on a 700 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 2048 points in F2 and 388 points in F1.

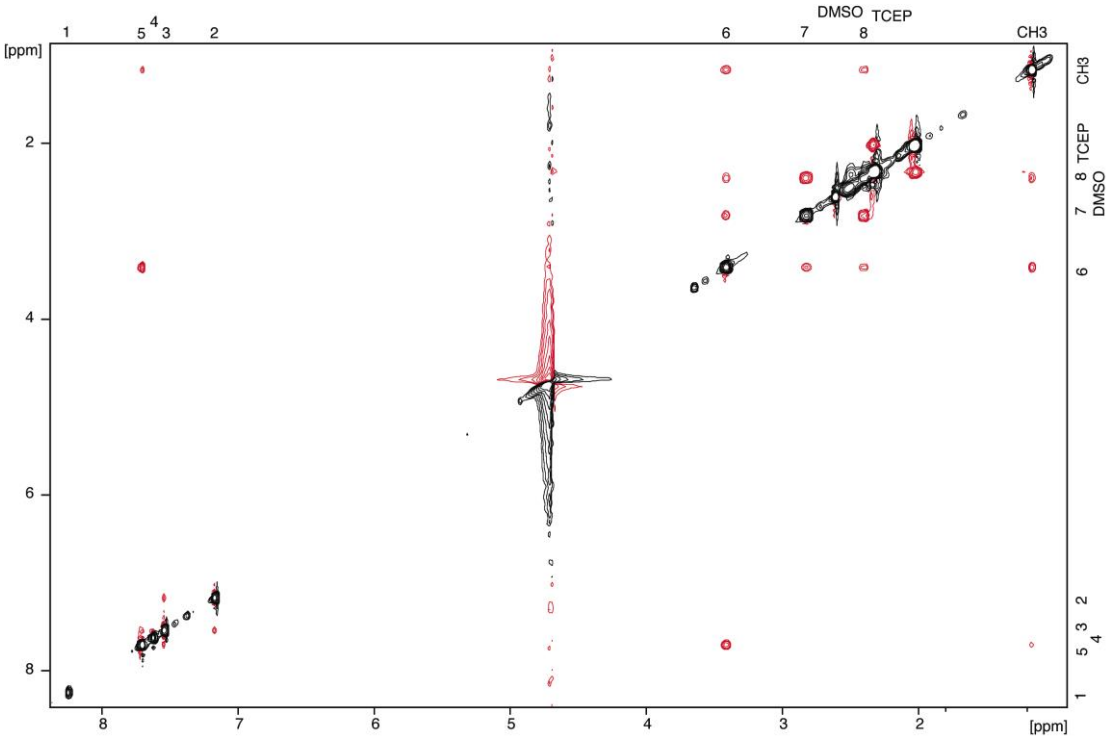


Figure S17c: NOESY spectrum of ligand 10 (1 mM) in buffer solution. In the free form of the ligand, no preferred conformation can be derived from the NOE. The mixing time was 800 ms on a 600 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 2048 points in F2 and 300 points in F1.

10 docking modes of ligand 10 in PKA were created and visually inspected. Surprisingly, none of the docked complex structures displayed the bound conformation of the ligand known from trNOE. This is not surprising since the ligand may adopt a conformation in the complex that is only little populated in the free form. Thus, the trNOE information is useful already when the docking is done. The bound conformation can be fixed during the docking process. Specifically, the docking software PLANTS allows to fix the distance between two protons within a certain range. The distances between the protons close in space, as interpreted based on the trNOE spectrum shown in Fig. S18, were restrained to the following distance ranges: H1-H4: 1-3 Å, H3-H5: 1-3 Å, and H5-CH3: 1-5 Å. These distance ranges were chosen, since a rotation of the torsion angle by 180° from the expected conformation resulted in a distance larger than 3 Å or 5 Å, respectively.

By contrast to the first approach in which all conformations of ligand 10 were taken into account, now only structures that fulfill the trNOE restraints were taken into account for the STD and INPHARMA evaluation (with ligand 2).

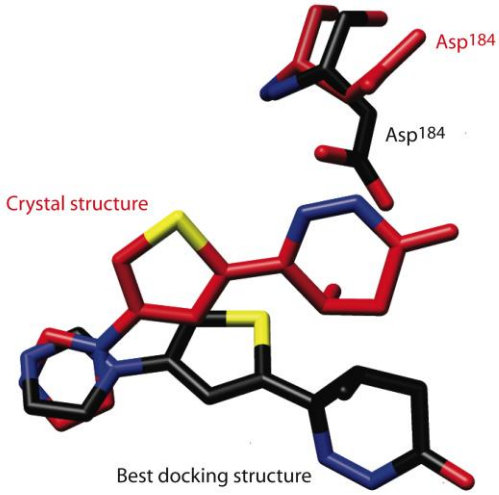


Figure S18: Best docking model of ligand 10 by R_{STI} in black and the crystal structure in red. Ligand orientation is correct, but the solvent exposed part of the ligand does not well reproduce the crystal structure. It can be seen that the residue Asp^{184} is rotated in the crystal structure and allows the ligand to adopt a binding mode which is restricted in the rigid protein structure that is used during the docking.

When the trNOE compatible structure was run in an unrestrained MD simulation for 1 ns the residue Asp^{184} changed its conformation in the first 3 ps and ligand 10 was allowed to adopt the correct orientation as can be seen in Fig. S20. The best binding mode scored with R_{STI} = 0.85 ($R_{INPHARMA}$: 0.77, R_{STD}: 0.84, R_{trNOE}: 0.93). The RMSD to the crystal structure is 1.12 Å.

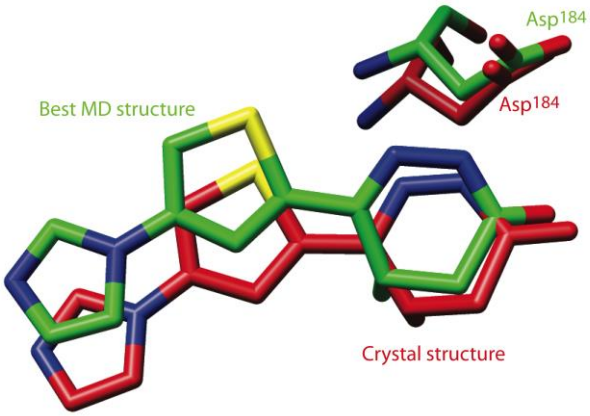


Figure 19: Best MD structure of ligand 10 by R_{STI} in green and the crystal structure in red. The residue Asp^{184} opens during the MD simulations and gives way for the correct binding mode, which scores significantly higher than the original docking structure.

The trNOE optimizes the binding mode substantially with the R_{STI} rising from 0.29 to 0.85. The conclusion is, that the conformation of a ligand with freely rotatable bonds can be fixed by trNOE during the docking. Then, the trNOE needs to be scored first to take only structures into account, that fulfill the trNOE restraints. These binding modes can than be further scored by STD and INPHARMA which opens up binding pockets that are otherwise not accessible

Crystal structures

Inhibitor complexes of the recombinant bovine Cα catalytic subunit of cAMP-dependent protein kinase with ligand were obtained by soaking the crystals for 24-48 hours in pre-equilibrated drops containing the crystallization conditions and 10 mM L3 or L4. The crystals were cryo-protected by the addition of 30% glycerol and frozen in liquid nitrogen. Diffraction data were collected at ESRF ID23-1. The datasets were processed, the structure was solved and initially refined with an internally developed pipeline using autoPROC[1], XDS[2] and autoBUSTER[3]. 5-7% of the reflections were set aside for cross-validation. Interactive model building and refinement was carried out using COOT[4] and autoBUSTER yielding a final R/R_{free} of 17.6%/23.1% for the PKA:L4 and 17.3%/19.7% for the PKA:L3 complex (Table S2). Coordinates and structure factors are available at the protein data base (PDB) under entry codes 4EXF (L3) and 4EXD (L4).

Table T1 Data collection and refinement statistics

	PKA:L4	PKA:L3

Data collection

Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, b, c (Å)	73.1, 75.4, 90.0	72.8, 75.1, 79.9
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	75-2.11(2.22-2.11)	20-2.00(2.08-2.00)
Rsym	9.8(48.5)	5.9(38.2)
I/σI	15.0(3.9)	21.1(4.7)
Completeness (%)	99.9(100)	99.7(99.2)
Multiplicity	5.5(5.5)	5.4(5.3)

Refinement

Resolution (Å)	54-2.11	19-2.00
No. reflections (work/free)	26050/1285	28054/2116
Rwork/Rfree	0.176/0.231	0.173/0.197
No. atoms		
Protein	2943	2931
Water	343	382
Ligand	17	10
Mean B-value (Å²)	26.9	31.0
R.m.s deviations		
Bond lengths (Å)	0.01	0.01
Bond angles (°)	1.05	1.04
Ramachandran analysis (%)		
favoured regions	97.1	97.7

allowed regions	2.3	1.7
outlier	0.6	0.6

*Highest resolution shell is shown in parenthesis.

Table T2a The table lists the structures of the ligands 1-4 obtained from docking to the crystal structure 3DNE with PLANTS and gives the RMSD to the respective crystal structure. Highlighted in colours is the best ranked structure from every scoring function. PLANTS applies the docking scoring function ChemPLP, which is also shown in the main text. The docking modes were alternatively scored with the scoring functions available from GOLD, namely ChemScore, Astex Statistical Potential (ASP) and GOLDScore. Please note that for ChemPLP, the smallest numbers are the best structures, while for the GOLD scoring functions the highest are the best. Before scoring, GOLD is performing an optimization of the torsion angles. For comparison, the STD Score is given. Hereby the largest numbers represent the best scores.

Name	RMSD to X-ray	ChemPLP	ChemScore	ASP	GOLDScore	R_{STD}
Ligand 1						
1	0.41	-84.5	33.8	28.27	45.78	0.575
2	2.56	-52.6	21	19.56	28.58	0.6132
3	2.62	-62.9	25.26	22.97	24.5	0.5041
4	3.94	-54.6	18.97	20.49	25.24	0.5138
5	5.1	-73.1	26.89	26	32.09	-0.134
6	5.24	-52.7	19.62	14.65	24.44	**0.6553**
7	5.5	-59.8	25.56	21.21	32.66	-0.1334
8	5.53	-63.3	25.8	26.87	34.22	0.1338
9	5.59	-54.3	23.8	19.31	30.93	0.0145
10	6.18	-53.5	19.78	19.69	29.4	-0.4638
11	6.33	-58.6	20.48	20.03	29.99	-0.1102
12	6.99	-52.4	17.01	15.37	21.84	-0.2668
13	7.25	-54	20.26	16.15	17.66	-0.4884
14	7.34	-53.7	17.69	15.73	24.37	0.4476
15	7.38	-54.5	19.63	20.66	21.07	-0.5749
16	7.83	-57.6	18.15	15.36	26.02	0.1949
Ligand 2						
1	1.41	-66.3	24.53	26.06	39.99	**0.0624**
2	3.14	-58.3	16.96	16.57	32.47	-0.2566
3	5.2	-70.9	27.91	26.24	36.22	-0.1466

4	5.26	-69.4	25.02	24.19	35.66	-0.1466
5	5.8	-58	17.12	17.47	31.92	-0.6284
6	6.78	-55.8	15.42	14.04	24.9	-0.7075
7	7.68	-61.5	17.33	13.16	30.25	-0.6575
Ligand 3						
1	0.96	-47	18.97	14.35	27.32	0.7036
2	2.47	-44.7	18.21	16.11	23.25	0.6329
3	2.85	-51.6	22.28	16.08	30.81	-0.3923
4	2.86	-55.4	24.25	21.5	35.89	**0.7304**
5	3.36	-60.5	25.53	20.42	31.26	0.4096
6	3.42	-42.8	16.78	15.12	28.7	-0.0246
7	3.74	-47.7	18.16	16.33	25.66	0.2344
8	3.78	-57.9	22.29	20.61	29.33	0.4731
9	4.4	-57.5	23.8	17.02	29.45	-0.4377
10	4.55	-43.4	16.99	14.55	19.32	-0.6908
11	4.55	-43	18.01	14.36	13.62	0.3948
12	4.59	-51.5	21.45	15.52	28.51	0.0048
13	5.44	-44.5	15.82	13.1	18.01	0.3694
14	5.54	-49.3	18.55	15.61	25.17	0.773
15	5.74	-44.1	16.21	12.41	12.73	0.2122
16	6.57	-47.7	16.76	11.53	21.5	0.2785
17	7.26	-44.5	13.63	13.1	20.69	-0.708
18	7.5	-43	14.36	12.56	20.65	-0.0272
19	7.83	-43.1	14.61	11.05	18.7	-0.1609
Ligand 4						
1	0.15	-64.3	16.7	27.5	40.39	**0.8751**
2	2.07	-56.1	11	24.43	30.65	-0.981
3	2.26	-57.8	11.11	23.94	28.63	0.8202
4	2.81	-54.8	10.95	21.98	20.48	-0.968
5	2.92	-58.8	14.4	25.6	37	-0.999
6	3.02	-56.4	11.19	22.12	38.04	-0.5229
7	3.85	-66	11.97	23.66	36.58	-0.942
8	4	-66.7	10.36	23.8	29.1	-0.9978
9	4.57	-58.5	10.68	21.59	34.51	-0.4117
10	5.92	-58.7	18.32	22.29	33.03	-0.992
11	5.96	-58.3	12.66	21.58	33.41	-0.9917
12	5.99	-62.5	17.25	24.3	29.35	-0.9891

13	6.01	-54.7	12.51	21.35	28.46	-0.7529
14	6.1	-63.8	17.09	25.5	34.44	-0.9905
15	6.29	-59.1	17.76	25.94	37.76	-0.9781
16	6.44	-55.9	15.94	26.62	35.21	-0.9295
17	6.94	-58.7	15.1	23.05	35.51	-0.9674
18	7.14	-57.4	16.67	24.11	29.15	-0.9995
19	7.18	-54.5	18.55	20.38	31.15	-0.8712
20	7.22	-61.4	17.43	22.66	35.46	-0.9997
21	7.26	-56.1	16.32	19.17	28.91	-0.7941
22	7.34	-56	12.65	20.05	28.97	0.1045
23	7.4	-59.7	16.58	22.82	32.46	-0.9474
24	7.43	-60.6	11.61	20.45	33.37	-0.9741
25	7.46	-55.5	10.89	20.05	29.37	-0.7615
26	7.6	-62.6	12.18	19.4	33.01	-0.9584
27	7.93	-54.5	13.71	18.14	26.73	-0.9698

Table T2b The tables list the R_{STI} values for the ligand combinations 1&2 and 3&4. The optimal combinations are highlighted.

Ligand 1 Name	Ligand 2 Name 1	2	3	4	5	6	7
1	**0.83**	0.54	0.41	0.41	0.48	0.22	0.18
2	0.82	0.55	0.41	0.32	0.43	0.27	0.11
3	0.32	0.12	-0.37	-0.16	-0.35	-0.45	0.05
4	0.49	0.35	0.02	0.13	-0.01	-0.32	0.01
5	0.55	0.29	-0.40	0.23	0.21	0.16	0.14
6	0.56	0.44	0.16	0.16	0.12	0.42	0.27
7	0.69	0.40	0.30	0.09	0.31	0.30	-0.06
8	0.40	0.08	-0.07	-0.19	-0.12	0.18	0.26
9	0.41	0.37	0.05	0.05	0.19	-0.05	0.07
10	0.50	0.26	-0.20	-0.33	-0.08	-0.24	0.19
11	0.24	0.17	-0.21	-0.21	-0.19	-0.04	0.40
12	0.72	0.54	0.16	0.12	0.26	0.13	0.17
13	0.11	0.00	-0.41	-0.46	-0.33	-0.26	-0.13
14	0.20	0.43	-0.34	-0.27	-0.11	0.21	0.15
15	0.78	0.52	0.38	0.43	0.36	0.01	-0.29
16	0.34	0.09	-0.12	-0.12	-0.39	-0.12	0.13

Ligand 4 Name	Ligand 3 Name 1	2	3	4	5	6	7
1	**0.81**	0.38	0.12	0.40	0.30	0.64	0.28
2	0.22	0.35	-0.22	0.11	0.06	-0.05	0.11
3	0.53	0.33	0.08	0.40	0.27	0.28	0.47
4	0.46	0.37	0.16	0.50	0.27	0.16	0.37
5	0.12	0.09	-0.05	0.18	0.12	-0.14	0.17
6	0.34	0.21	-0.04	0.23	0.11	0.08	0.23
7	-0.01	0.28	-0.41	-0.10	0.01	-0.10	-0.12
8	-0.09	-0.03	-0.40	-0.07	-0.08	-0.23	-0.06
9	0.35	0.35	-0.08	0.27	0.18	-0.02	0.24
10	0.01	0.20	-0.39	-0.11	-0.11	-0.18	-0.13
11	0.24	0.09	-0.32	-0.02	-0.05	-0.13	0.00
12	0.14	0.20	-0.25	0.08	0.09	-0.06	0.10

13	0.24	0.55	-0.23	0.13	0.12	-0.17	0.15
14	0.00	0.37	-0.20	0.00	0.00	-0.22	-0.05
15	-0.01	0.01	-0.22	-0.04	-0.08	-0.26	-0.06
16	0.19	0.27	-0.14	0.08	0.11	-0.08	0.02
17	-0.02	-0.06	-0.42	-0.16	-0.22	-0.26	-0.12
18	0.01	-0.02	-0.32	-0.06	-0.10	-0.22	-0.07
19	0.02	-0.12	-0.28	-0.07	0.05	-0.19	-0.08
20	-0.01	-0.01	-0.33	-0.11	-0.11	-0.24	-0.11
21	0.02	-0.01	-0.28	-0.05	-0.03	-0.23	-0.03
22	0.37	0.35	-0.05	0.28	0.28	0.13	0.28
23	0.07	0.03	-0.20	0.02	0.02	-0.20	0.07
24	0.00	0.36	-0.37	-0.11	-0.05	0.10	-0.11
25	0.22	0.26	-0.15	0.22	0.22	-0.08	0.09
26	0.42	0.53	0.09	0.43	0.43	0.22	0.42
27	0.48	0.45	0.06	0.39	0.42	0.30	0.39

Ligand 4 Name	Ligand 3 Name 8	9	10	11	12	13	14
1	0.40	0.12	0.71	0.27	0.20	0.40	0.38
2	0.08	-0.16	0.26	-0.06	-0.01	0.24	0.13
3	0.39	0.13	0.29	0.27	0.28	0.44	0.36
4	0.38	0.19	0.31	0.24	0.29	0.39	0.39
5	0.20	-0.12	0.14	0.05	0.15	0.18	0.20
6	0.23	-0.04	0.09	0.06	0.15	0.25	0.25
7	-0.09	-0.35	0.00	-0.07	-0.09	-0.06	-0.05
8	-0.06	-0.34	0.07	-0.14	-0.19	-0.06	-0.03
9	0.29	0.12	0.21	0.18	0.11	0.26	0.30
10	0.01	-0.32	-0.15	-0.20	0.02	-0.04	0.00
11	0.10	-0.23	-0.03	-0.08	0.36	0.03	0.13
12	0.18	-0.19	0.03	-0.07	-0.04	0.10	0.13
13	0.19	-0.18	0.09	0.04	0.11	0.17	0.18
14	0.04	-0.30	-0.07	-0.19	0.04	-0.03	0.11
15	-0.08	-0.35	-0.14	-0.25	-0.05	-0.06	-0.04
16	0.20	-0.17	0.05	0.04	0.24	0.06	0.20
17	0.10	-0.28	-0.14	-0.21	0.25	-0.07	-0.02
18	-0.03	-0.33	-0.09	-0.16	-0.14	-0.06	-0.02
19	0.31	-0.26	-0.07	-0.10	-0.01	-0.05	0.03

20	-0.04	-0.31	-0.06	-0.18	-0.06	-0.09	0.03
21	0.02	-0.28	-0.08	-0.12	-0.08	-0.02	0.02
22	0.27	0.00	0.22	0.21	0.25	0.27	0.32
23	-0.02	-0.27	-0.03	-0.05	0.02	0.03	0.08
24	-0.02	-0.36	-0.03	-0.18	-0.17	-0.06	-0.05
25	0.49	-0.10	0.14	0.11	0.17	0.11	0.36
26	0.42	0.14	0.37	0.33	0.27	0.42	0.46
27	0.47	0.11	0.61	0.25	0.37	0.42	0.46

Ligand 4 Name	Ligand 3 Name 15	16	17	18	19
1	0.36	0.18	0.22	0.10	0.35
2	0.15	-0.20	0.01	-0.19	-0.01
3	0.37	0.18	0.15	0.12	0.34
4	0.40	0.09	0.14	0.13	0.33
5	0.19	-0.15	-0.03	-0.17	0.08
6	0.21	0.01	-0.09	-0.02	0.19
7	-0.05	-0.34	-0.22	-0.38	-0.22
8	-0.03	-0.31	-0.28	-0.39	-0.14
9	0.32	-0.01	0.04	0.06	0.17
10	0.04	-0.33	-0.34	-0.31	-0.11
11	0.05	-0.24	-0.28	-0.17	-0.05
12	0.14	-0.10	-0.19	-0.25	-0.03
13	0.19	-0.14	-0.10	-0.17	0.06
14	0.03	-0.32	-0.36	-0.28	-0.15
15	-0.02	-0.40	-0.37	-0.34	-0.21
16	0.13	-0.19	-0.23	-0.12	-0.02
17	-0.05	-0.35	-0.41	-0.32	-0.08
18	-0.03	-0.36	-0.34	-0.29	-0.13
19	0.01	-0.28	-0.56	-0.23	-0.14
20	-0.04	-0.36	-0.36	-0.29	-0.13
21	0.03	-0.23	-0.29	-0.25	-0.07
22	0.31	-0.03	0.01	0.06	0.20
23	0.12	-0.26	-0.29	-0.21	0.02
24	0.00	-0.40	-0.29	-0.38	-0.20
25	0.30	-0.01	-0.12	-0.06	0.13
26	0.46	0.10	0.18	0.09	0.30

27	0.43	0.15	0.16	0.15	0.37

Part II: Glycogen phosphorylase (GP)

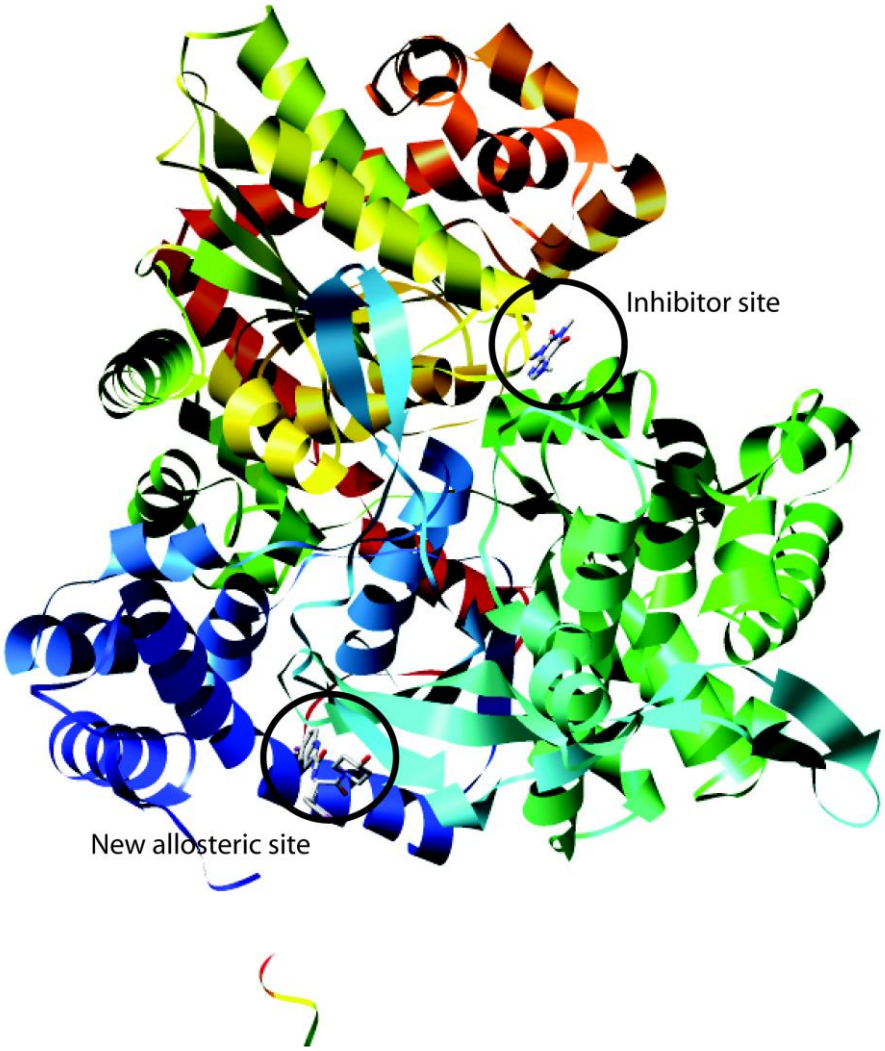


Figure S20: PDB crystal structure 1C8L of glycogen phosphorylase (GP) to illustrate the positions of the inhibitor site and the new allosteric site. At the inhibitor site, bound caffeine is shown. The PDB structure 1C50 of GP was superimposed with 1C8L to show the ligand CP320626 binding to the new allosteric site. For clarity, only PDB structure 1C8L is shown.

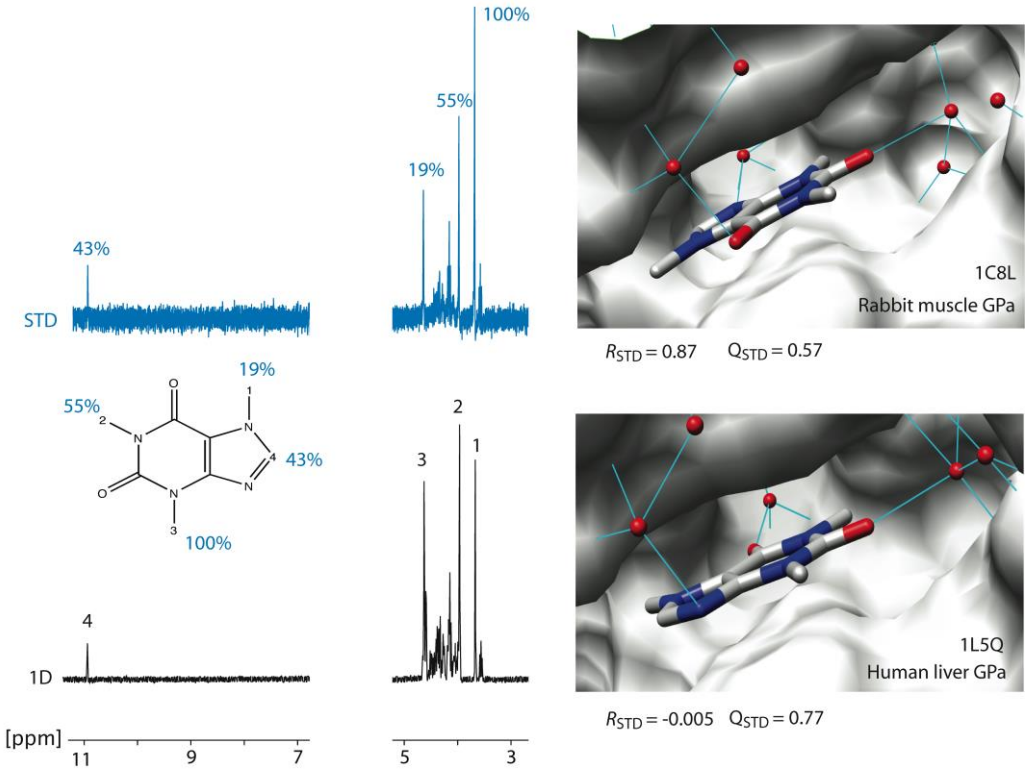


Figure S21: Experimental STD data of caffeine (400 µM). The normal 1D proton NMR spectrum in presence of 2 µM GP is shown in black, the STD spectrum in blue. STD saturation was applied for 2 s on 0.5 ppm on a 600 MHz spectrometer (temperature = 298 K). There are two different binding modes of caffeine to GP to be found in the PDB which are rotated 180° in comparison to each other.

These PDB structures, 1C8L from rabbit muscle GPa and 1L5Q from human liver GPa, are shown on the right side. The correlation between experimental and back-calculation of STD data R_{STD} and Q_{STD} gives a clear answer that the binding mode of rabbit muscle GPa is present in the applied study. Therefore 1C8L was chosen to give the coordinates of bound caffeine to GP.

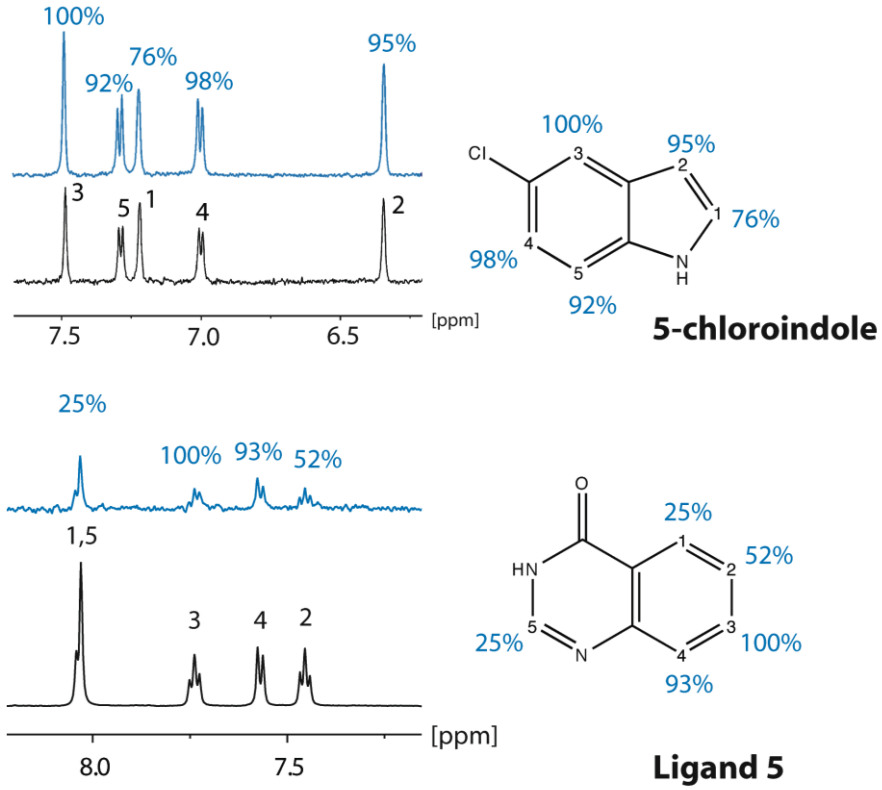
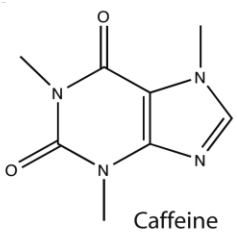
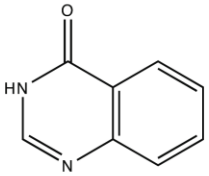


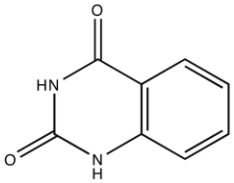
Figure S22: Experimental STD data of 5-chloroindole and ligand 5 (400 µM each). The normal 1D proton NMR spectrum in the presence of 2 µM GP is shown in black, the STD spectrum in blue. STD saturation was applied for 2 s on 0.5 ppm on a 600 MHz spectrometer (temperature = 298 K).



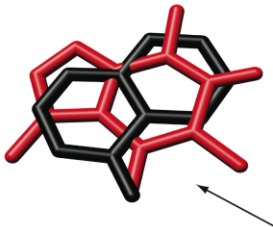
Caffeine

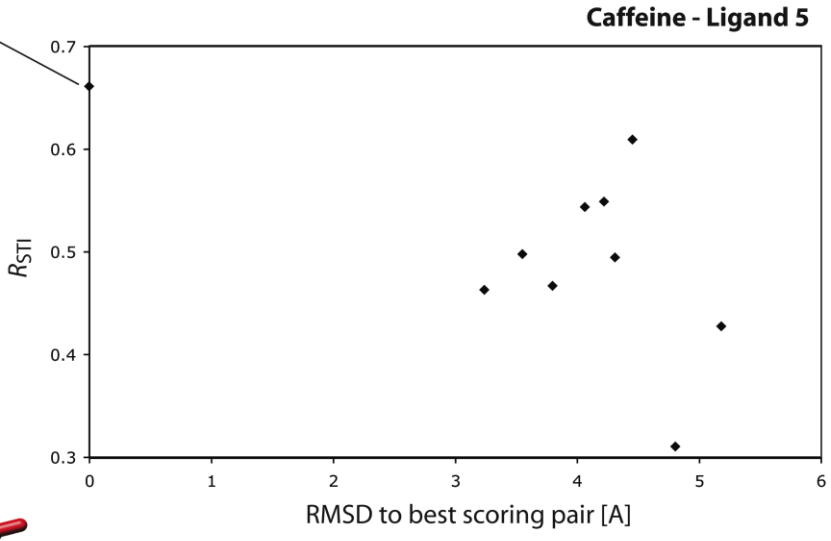


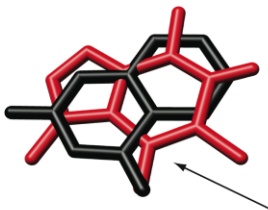
Ligand 5



Ligand 7







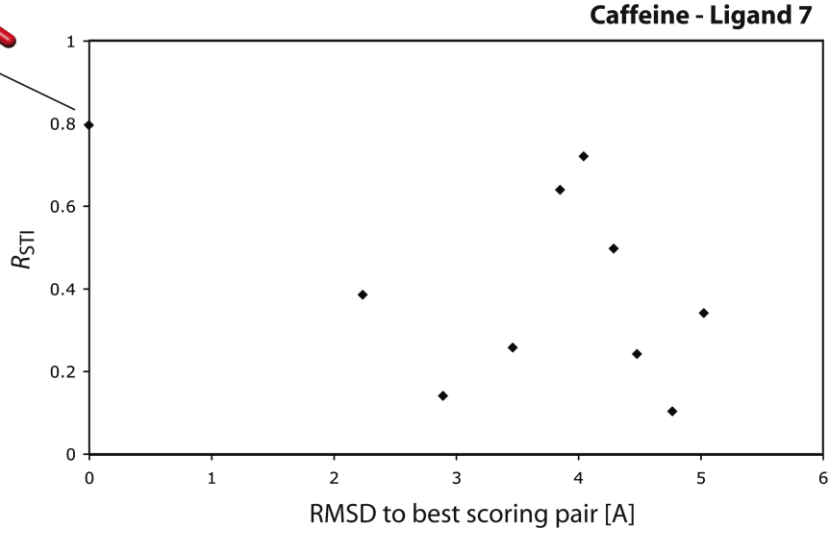


Figure S23: R_{STI} plotted against the RMSD of 10 docking structures of ligand 5 and 7, respectively, to the structure with the best R_{STI} score. Caffeine is fixed within the binding mode of crystal structure 1C8L. For ligands 5 and 7, a similar binding mode is found where the oxygen forms a common pharmacophore with the oxygen of caffeine. The ligand is placed between two phenylalanine residues and therefore only rotations by 180° are allowed by the binding site. This explains the large difference in RMSD between the best scoring structure and the other proposed binding modes.

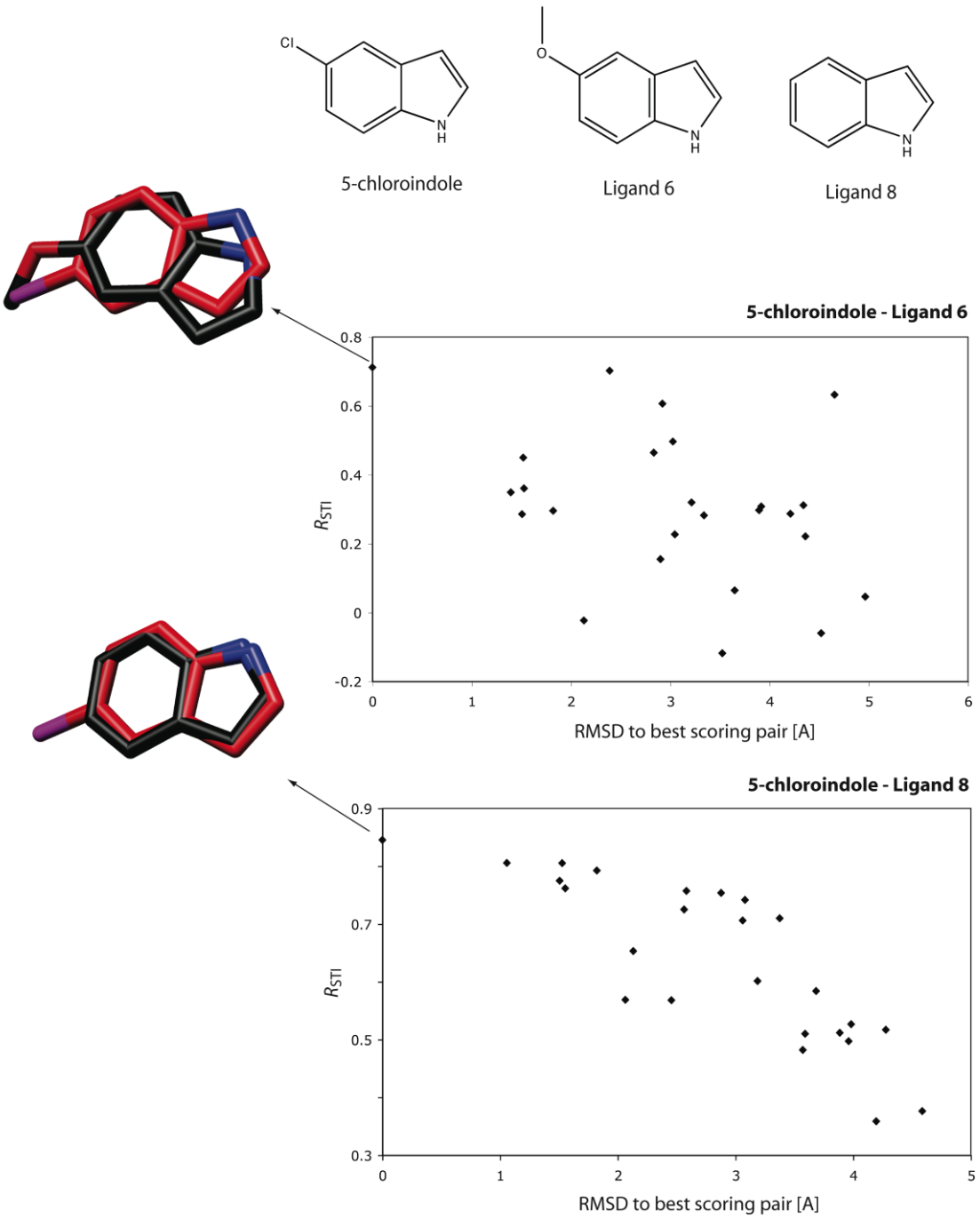


Figure S24: R_{STI} plotted against the average RMSD of 5 docking structures of chlor-indole and 5 docking structures of 6 and 8, respectively, to the structure with the best R_{STI} score. For ligands 6 and 8, a similar binding mode is found were the nitrogens superimposed to the nitrogen of 5-chloroindole. Additionally, the methyl group of ligand 6 is a bulky moiety, superimposed with the chlorine of chlor-indole. The differentiation of R_{STI} is not as pronounced as in the other examples. Still, the result seems reasonable since the strongest INPHARMA peaks of the ligand 6 methyl group are observed to the protons of chlor-indole in the ortho positions to the chlorine atom.

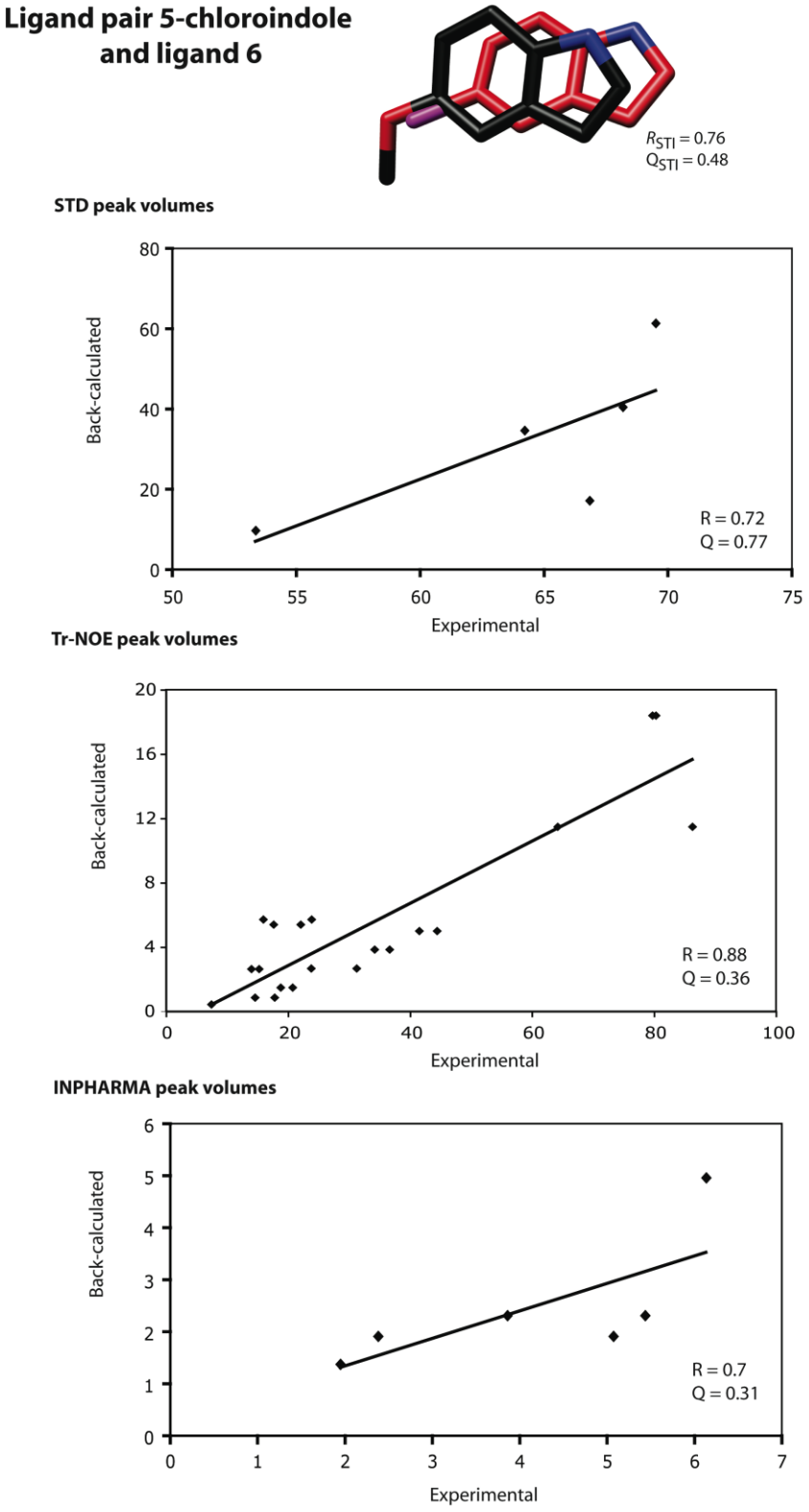


Figure S25: Correlation of the experimental and back-calculated peak volumes of ligands 5-chloroindole and 6 with the best R_{STI} score, refined during an MD simulation against STD, trNOE and INPHARMA peak volumes, respectively.

Part III: Soluble epoxide hydrolase (sEH)

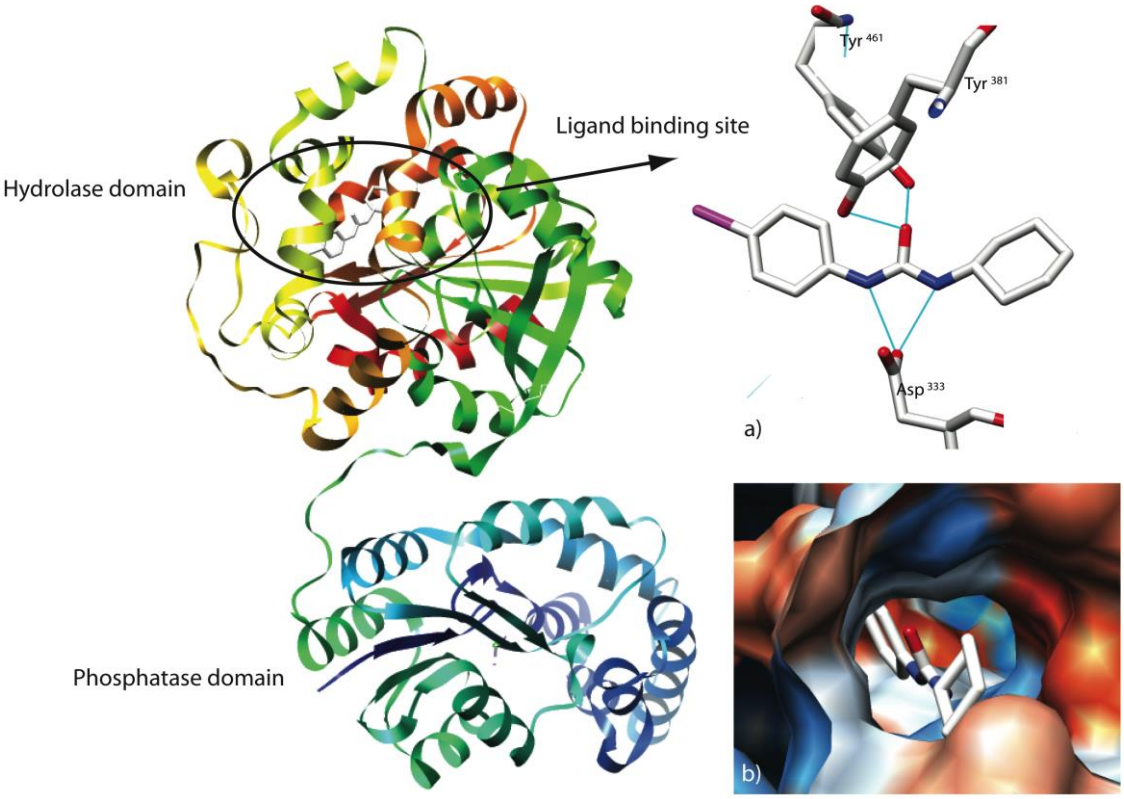


Figure S26: PDB structure 1VJ5 of soluble epoxide hydrolase (sEH). The hydrolase domain with the bound ligand and the phosphatase domain is shown. In the presented study only the hydrolase domain was expressed during sEH production as described below. A more detailed look on the binding site is taken on the right side: a) The urea center of the ligand is anchored between two tyrosines and an aspartic acid. This anchor is in the center of a tunnel which the ligand can enter from both sides. The binding sites around the central anchor are termed here Met^{337} side (cyclohexane substituent) and Met^{418} side (aromatic substitutent), whereas in b) The ligand is seen from the Met^{337} side.

Expression and purification of soluble epoxide hydrolase (sEH)

The human soluble epoxide hydrolase (sEH) consists of two domains and the x-ray structure reveals[5] that the protein forms a dimer with a domain-swapped architecture. The N-terminal domain has a phosphatase activity, while the epoxide hydrolase activity is within the C-terminal domain[5,6]. Whether the physical connection of these two domains with different catalytic activities has any physiological relevance is unknown.

In order to focus on the epoxide hydrolase activity, a corresponding construct for expression of the epoxide hydrolase domain was designed. This construct comprises amino acids 230-555 of sEH and expression was achived in reasonable amounts in *E.coli*.

The coding sequence for amino acids 230-555 of human sEH (GeneBank accession: NM_001979) with eight N-terminal consecutive histidine residues followed by a Tobacco etch virus (Tev)-protease cleavage site was cloned into a pET16b vector (Novagen, Merck, Darmstadt, Germany). This construct was transformed in *E.coli* Bl21(DE3) RIPL (Stratagene, La Jolla, CA, USA). Expression was done using an autoinducing medium[7] allowing the cells to grow for 18 hours at 30°C. Cells were harvested by centrifugation and resuspended in 500 mM NaCl, 50 mM Tris, 5 mM beta-mercaptoethanol, pH 8.0, supplemented with 2U/mL Benzonase (Merck, Darmstadt, Germany) and complete EDTA-free protease inhibitor tablets (Roche Diagnostics, Mannheim, Germany) The resuspended cells were lysed with the aid of a cell disprution system (Z-Plus model, Constant Systems Ltd, U.K.) at 1.8 kbar. After disruption the lysate was centrifuged for 90 min at 13.700 *g* at 4°C. Capture was done on a HisTrap column (GE Healthcare, Freiburg, Germany) equilibrated with 500 mM NaCl, 50 mM Tris, 5 mM imidazol, pH 8.0. Protein was eluted from the column by applying a linear gradient to 500 mM imidazol in the same buffer. Protein was digested with Tev-protease at 16°C over night in 100 mM NaCl, 50 mM Tris, pH 7.5, 2 mM DTT. The digest was passed a second time over the HisTrap column and the flowthrough, containing the cleaved protein, was collected. Further purification was done by anion exchange chromatography using a HiTrap Q column (GE Healthcare). The column was equilibrated with 50 mM sodium phosphate, pH 8.0, 2 mM DTT, and the protein was eluted with a gradient to 1.5 M NaCl. Further polishing was done by gelfiltration using a Superdex 200 (GE Healthcare) column equilibrated in 100 mM sodium phosphate, pH 7.4, 5 mM DTT.

Crystal structures of sEH

The sEH C-terminal domain (T230-M555) was crystallized using the hanging drop method: 1 µl of protein solution, containing 8 mg/ml sEH in 100mM Na_2HPO_4/NaH_2PO_4 (pH 7.4), 5mM DTT and 15% Glycerol was mixed with 1 µl reservoir solution, containing 100 mM Tris-HCL (pH 8.3), 26% PEG4000 and 200 mM Li_2SO_4, and set to equilibrate at 20°C. Large, square crystals appeared in about two weeks. **C9** and **A4** were soaked into native sEH crystals by transferring the crystals to a drop of 9 µl reservoir solution with 1 µl of a 100 mM inhibitor solution in DMSO. For **A8**, the inhibitor was added to the protein solution prior to crystallization. For cryoprotection, 20% glycerol was added to the soaking solution and the crystal was picked with a small nylon loop and flash frozen in liquid nitrogen. For **C9** and **A4**, data were measured on beamline PX-III at the Swiss Light Source (SLS) in Villigen, Switzerland. For **A8**, data were measured at Sanofi on a Mar345DTB image plate detector, mounted on a Rigaku Micromax007HF X-ray generator. Data were processed with XDS[2] and scaled with scala[8,9] as implemented in the autoproc procedure[1]. The structures were solved using a previously solved sEH structure as a starting model. The initial model was obtained by molecular replacement with Phaser[10] using a truncated structure of the full-length human sEH ([5]; pdb code 1S8O) as a starting model. Model building was done with Coot[4] and refinement was done with Buster[3].

Coordinates and structure factors are available at the protein data base (PDB) under entry codes 4C4X (C9), 4C4Y (A4) and 4C4Z (A8).

Crystallographic studies on sEH

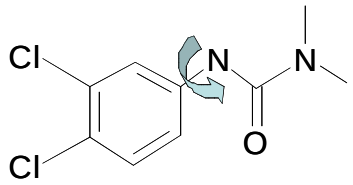
Crystals of sEH diffracted to 2.05-2.41 Å (Table T4), which allows the calculation of reasonably well-defined electron density maps. However, the pseudo-symmetric form of the inhibitors and their somewhat flexible binding mode (see below) did not allow a straightforward fitting of the inhibitors. Basically, two binding poses were possible, rotated 180° around the central urea moiety. To solve this problem, all possibilities were fitted and refined, i.e. a mixed binding mode, containing both conformations and two single binding poses with each one of the two possibilities. For the inhibitors, group occupancies were refined, where for the mixed binding poses the sum of the occupancies was constrained to be one. To judge the quality of the fits, the refined occupancies of the inhibitors were considered as well as the real-space correlation factor between the observed and calculated electron density of the inhibitors and the residual electron densities. Positive residual density indicates regions where, according to the experimental data, something should be fitted, which is missing from the model. Negative residual density indicates regions where something has been fitted, which should not be there according to the experimental data.

The results are given in Table T3 and Figure S27. The binding mode of the three inhibitors is similar to the binding mode of other sEH inhitors with the carbonyl oxygen of the central urea moiety interacting with Tyr^{383} and Tyr^{466}, and the nitrogens with Asp^{335}. In the A-conformation, the aromatic part of the inhibitor is bound in a hydrophobic pocket, lined by Phe^{267}, Tyr^{383}, Leu^{408}, Met^{419}, Val^{498},

His^{524} and Trp525, and the aliphatic part is bound in a pocket lined by Trp^{336}, Met^{339}, Thr^{360}, Tyr^{383}, Gln^{384}, Tyr^{466} and Leu^{499}. In the B-conformation, the aromatic and aliphatic parts of the inhibitors have swapped position.

With all three inhibitors, the A-conformation is the major, if not unique conformation, as can be deduced from the refined occupancies: 0.55 to 0.76 for the A-conformation versus 0.24 to 0.45 for the B-conformation; the higher real space correlations: 0.81 to 0.85 for the A-conformation versus 0.69 to 0.79 for the B-conformation; and from the lower residual densities for the A-conformation. Below, we will discuss the binding poses of the individual inhibitors.

C9: The most striking feature for this inhibitor is the absence of strong electron density for both electron-dense chloride ions (see Fig. S27). The most likely explanation for this phenomenon is the presence of conformational variability along the bond between the aromatic ring and the urea-nitrogen:



The result will be that the meta-chloride will occupy a range of positions, resulting in a smearing-out of its electron density, which may become too weak to be visible in our electron density maps. Also, the strong residual density, present for both the A- and B-conformation and in both sEH-monomers when refined as single conformation, strongly suggests that for this inhibitor genuinely two conformations are present, which will further blur the electron density.

A4: This inhibitor is highly symmetric, meaning that the electron density distribution of the inhibitor in conformation A will overlap very significantly with the electron density distribution of conformation B, making refinement of individual occupancies very difficult. Looking at residual density, which is very sensitive to the local fit in the electron density (Fig. S27), we see green blobs of positive residual density both in the refinement with both conformations and in the refinement with conformation B, but not in the refinement with conformation A. This means that the experimental data is best explained by the model with conformation A only.

A8: This inhibitor is the most asymmetric of the three inhibitors discussed here, making refinement of the occupancies of the two conformations more feasible. Looking at the residual electron density maps, we see some extremely small spikes of positive residual density with both conformations, some positive residual density with conformation A, and a lot of residual positive density (partly no longer visible since the refinement program has placed waters in it) with conformation B. This means that the inhibitor is mainly present in conformation A, with some conformation B. This corresponds well with the refinement occupancies: 0.76 for conformation A and 0.24 for conformation B.

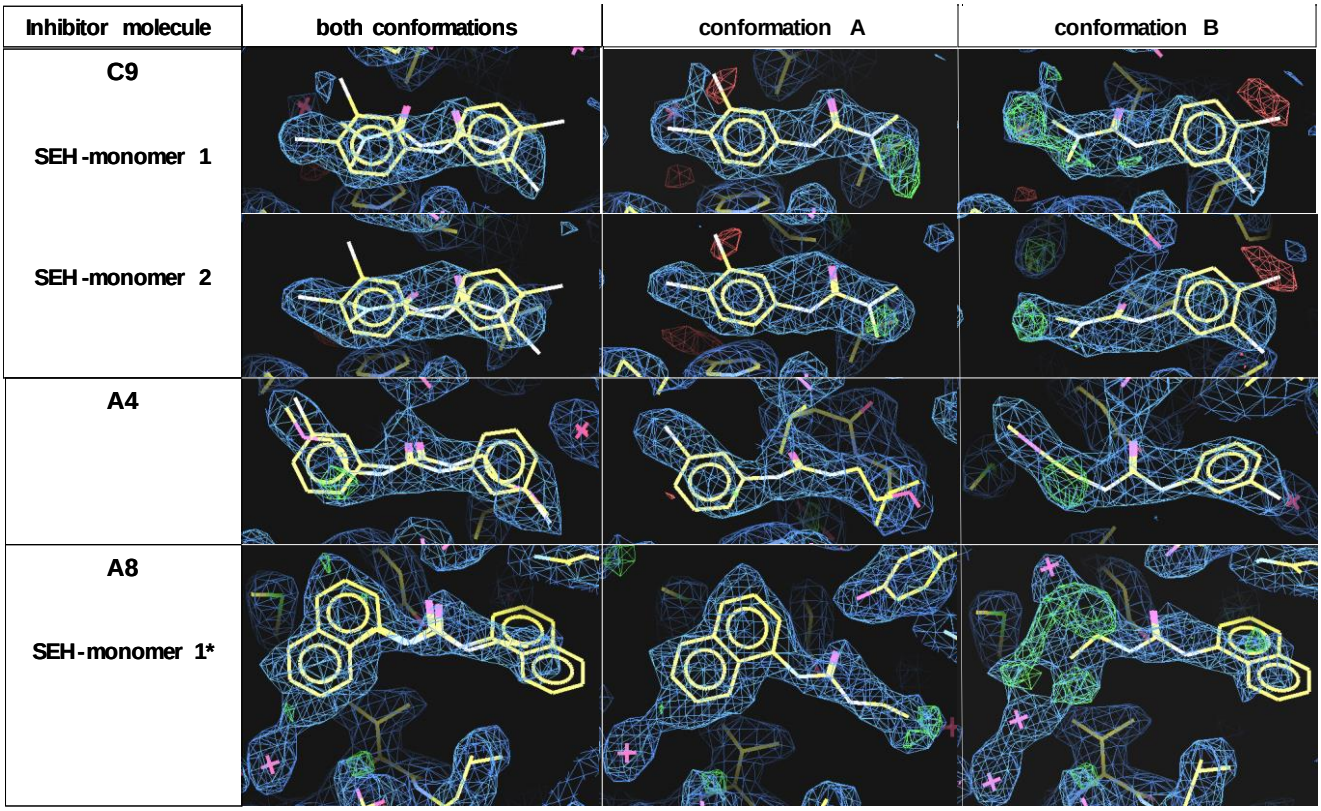


Figure S27: Electron density after refinement of the different binding poses. Blue: *2mFo-dFc* map, contoured at 1σ, green, red: *mFo-dFc* residual map, countered at +3σ and -3σ, respectively.

*for A8 monomer 2, no density whatsoever was present to fit the B-conformation, so only the A-conformation was fitted.

Table T3: Inhibitor statistics of the various binding modi.

	Mixed binding pose				Single binding pose			
Inhibitor molecule	**occ.A**	**occ.B**	**correlation**	**residuals**	**corr.A**	**residuals**	**corr.B**	**residuals**
C9-m1	0.60	0.40	0.85	0	0.81	+ −	0.75	++ −
C9-m2	0.60	0.40	0.85	0	0.81	+ −	0.73	+ −
A4	0.55	0.45	0.83	+	0.85	0	0.79	+
A8-m1	0.76	0.24	0.84	0	0.85	+	0.69	++
A8-m2*	0.99	-	0.86	0	0.86	0	-	-

Abbreviations and explanations: occ.A/B: refined group occupancy of the inhibitor in orientation A or B; residuals: residual density, +: positive blob >3σ, −:negative blob <-3σ; correlation/corr.A/B: real space correlation for the mixed binding pose or single orientation A or B; m1/2: sEH monomer 1 or 2 (for crystals forms with 2 sEH monomers in the asymmetric unit).

*for A8-m2, no density whatsoever was present to fit the B-conformation, so only the A-conformation was fitted.

Table T4 Crystallographic Data Collection and Refinement Statistics

	s**EH:C9**	**sEH:A4**	**sEH:A8**
Data collection			
space group	$P2_1$	$P2_122_1$	$P2_1$
cell dimensions			
a, b, c (Å)	47.69,	81.64	46.49
	79.87,	46.34	79.85
	88.61	80.29	87.22

α, β, γ (°)	90, 89.95, 90	90, 90, 90	90, 89.38, 90
resolution (Å)*	44.31-2.17(2.28-2.17)	57.24-2.41(2.54-2.41)	79.85-2.05 (2.16-2.05)
<I>/σ<I>	7.9(2.4)	15.0(4.6)	21.7(8.3)
completeness (%)	98.9(99.4)	98.9(97.8)	90.7(82.6)
redundancy	3.3(3.3)	5.6(5.7)	2.7(2.5)
Rmerge (%)	8.6(40.9)	7.5(32.7)	3.2(15.5)
unique reflections	35085	12191	36430
Refinement			
Conformation	both	A	A
Monomers in a.s.u.	2	1	2
Resolution (Å)	29.6-2.17	40.8-2.41	38.7-2.06
No. reflections (work/free)	32311/2512	11327/579	32998/2549
Rwork/Rfree	0.211/0.242	0.258/0.315	0.220/0.253
No. atoms			
Protein	5061	2535	5124
Water	445	139	602
Ligand	56†	15	32
Mean B-value (Å²)	38.6	37.0	27.2
R.m.s deviations			
Bond lengths (Å)	0.007	0.007	0.007
Bond angles (°)	0.94	0.93	0.92

*Highest resolution shell is shown in parenthesis.

†2 monomers and 2 conformations

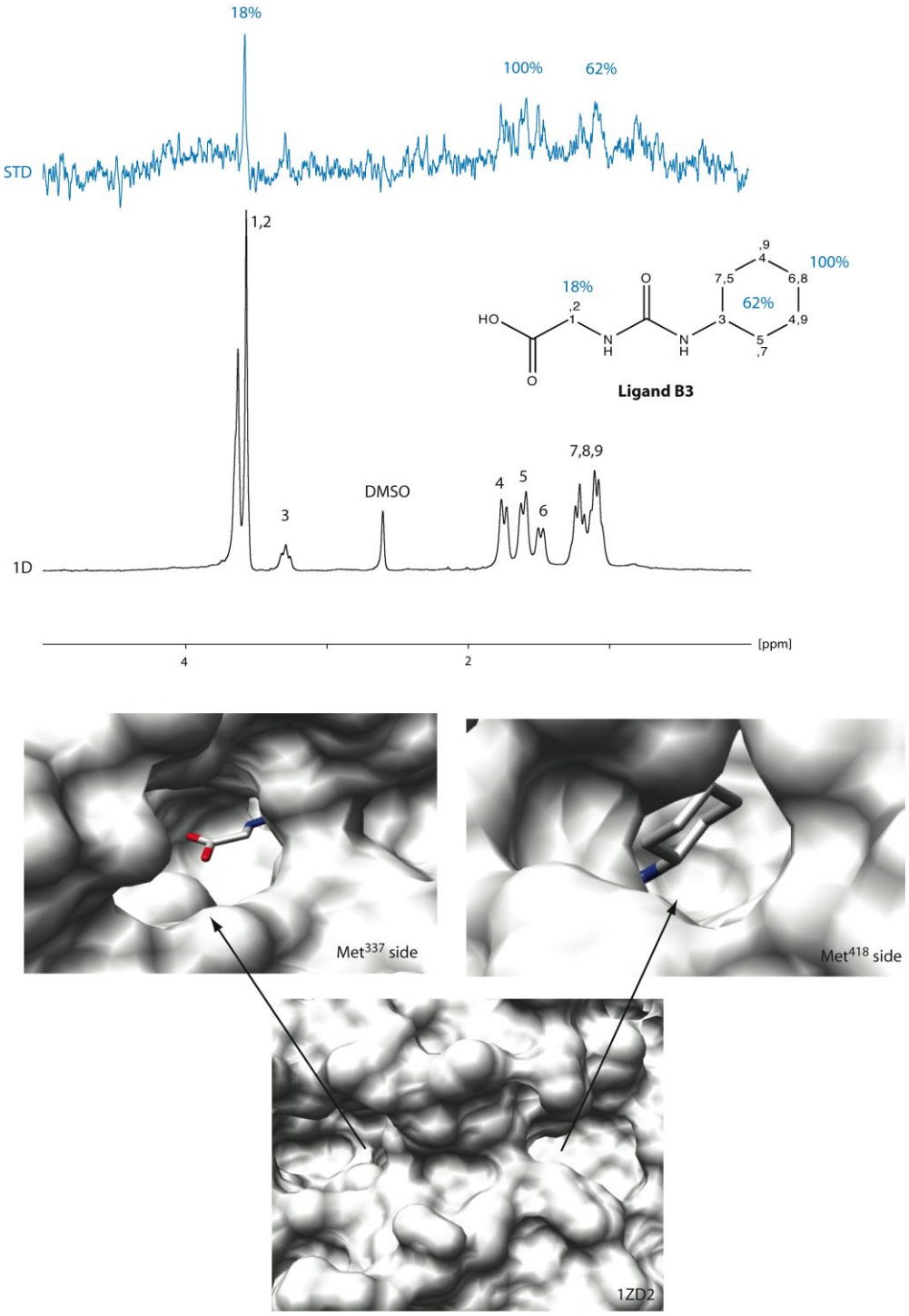


Figure S28: Experimental STD data of sEH binding ligand B3 (1 mM). The normal 1D proton NMR spectrum in the presence of 10 µM sEH is shown in black and the STD spectrum in blue. Saturation was applied for 8 s on 0.5 ppm on a 400 MHz spectrometer (number of scans = 8; temperature = 298 K; time domain = 16384 points). The ligand gave a weak STD signal, but the binding was too weak (IC_{50} > 500 µM) to obtain trNOE. The STD values fit very well to the crystal structure 1ZD2 (R_{STD} = 0.79; Q_{STD} = 0.53) of this ligand, where the cyclohexane substituent is oriented to the Met^{418} site and the carboxylic acid moiety to the Met^{337} site as illustrated in the figure.

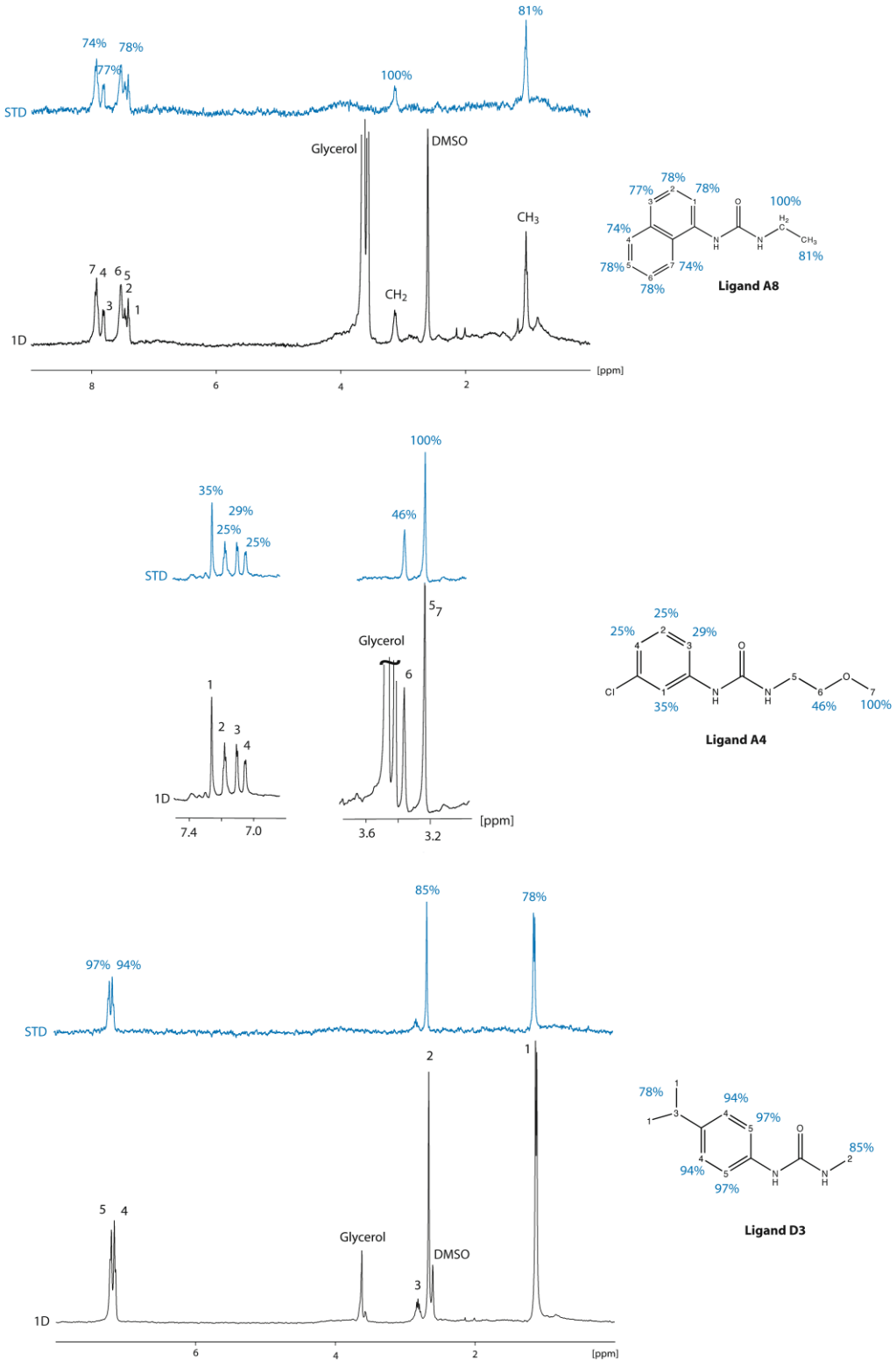


Figure S29: Experimental STD data of sEH binding ligands. The normal 1D proton NMR spectrum in the presence of 10 µM sEH is shown in black and the STD spectrum in blue. Ligand concentrations were 250 µM (A8), 1 mM (A4) and 1 mM (D3), respectively. Saturation was applied for 8 s on 0.5 ppm on a 400 MHz spectrometer (number of scans = 8; temperature = 298 K; time domain = 16384 points).

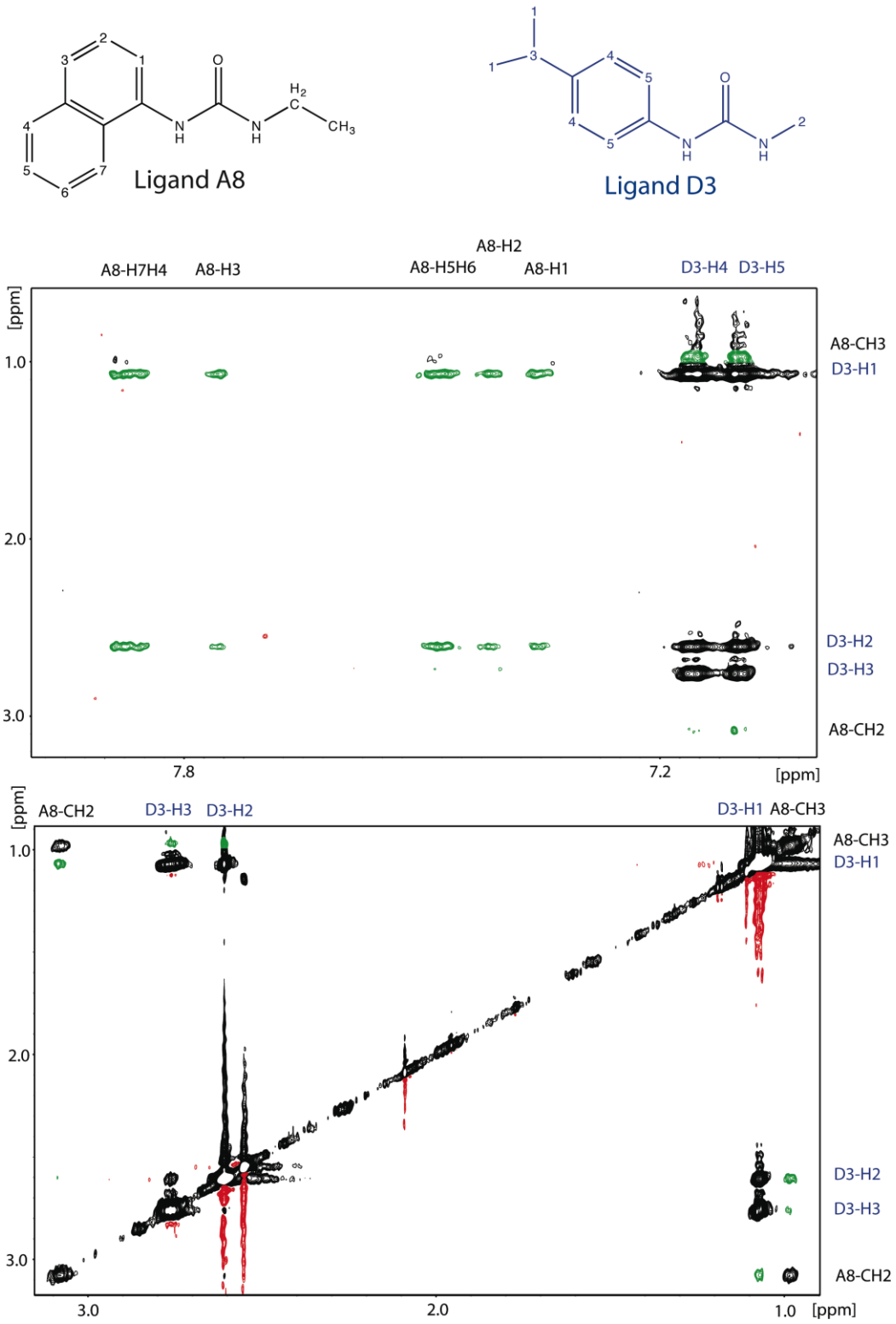


Figure S30: NOESY spectrum of ligand A8 (250 µM) and ligand D3 (500 µM) in the presence of sEH (40 µM). INPHARMA peaks chosen are marked in green. The mixing time was 800 ms on a 600 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 4096 points in F2 and 384 points in F1.

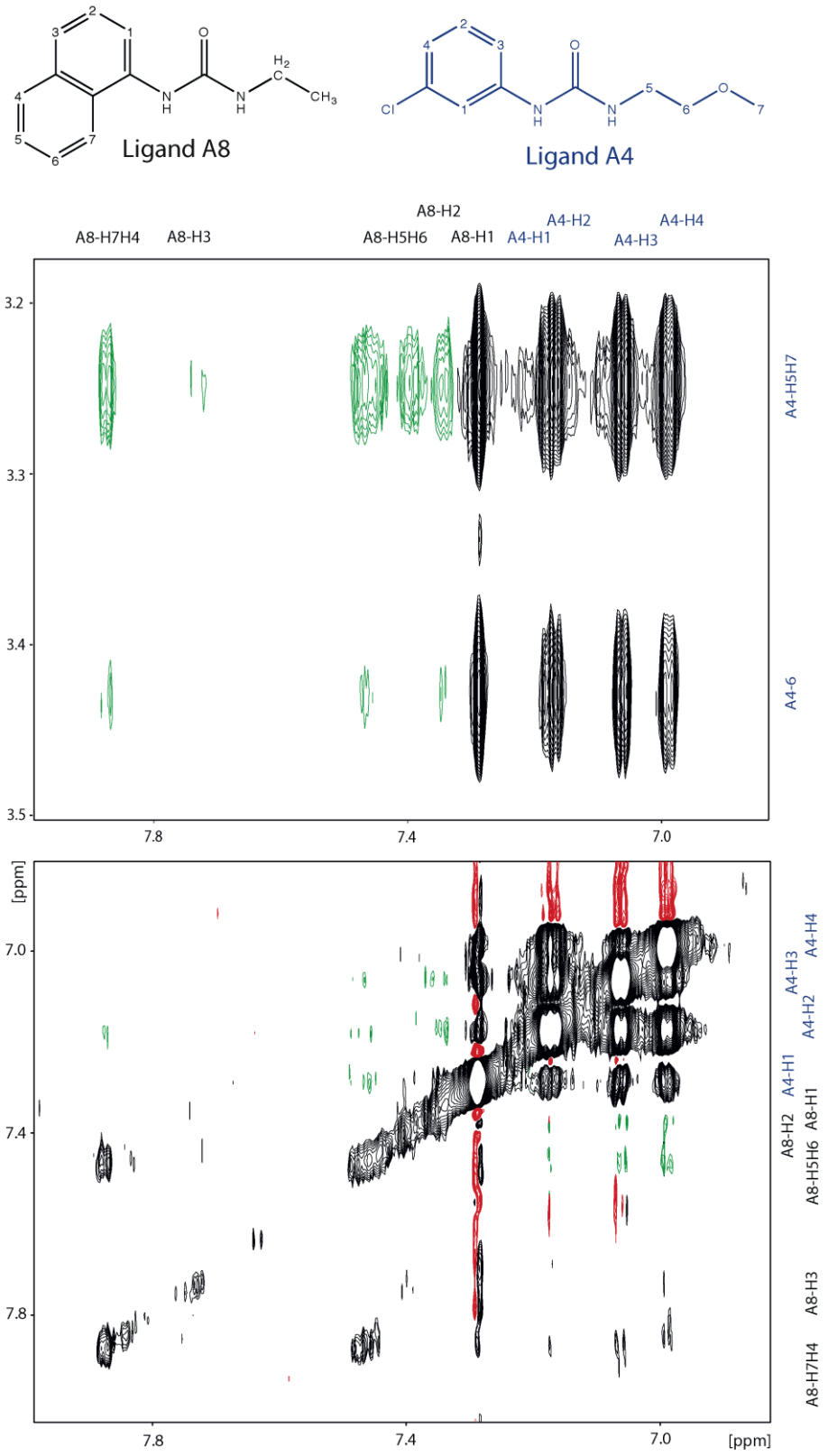


Figure S31: NOESY spectrum of ligand A8 (250 µM) and ligand A4 (750 µM) in the presence of sEH (30 µM). INPHARMA peaks chosen are marked in green. The mixing time was 800 ms on a 700 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 8192 points in F2 and 384 points in F1.

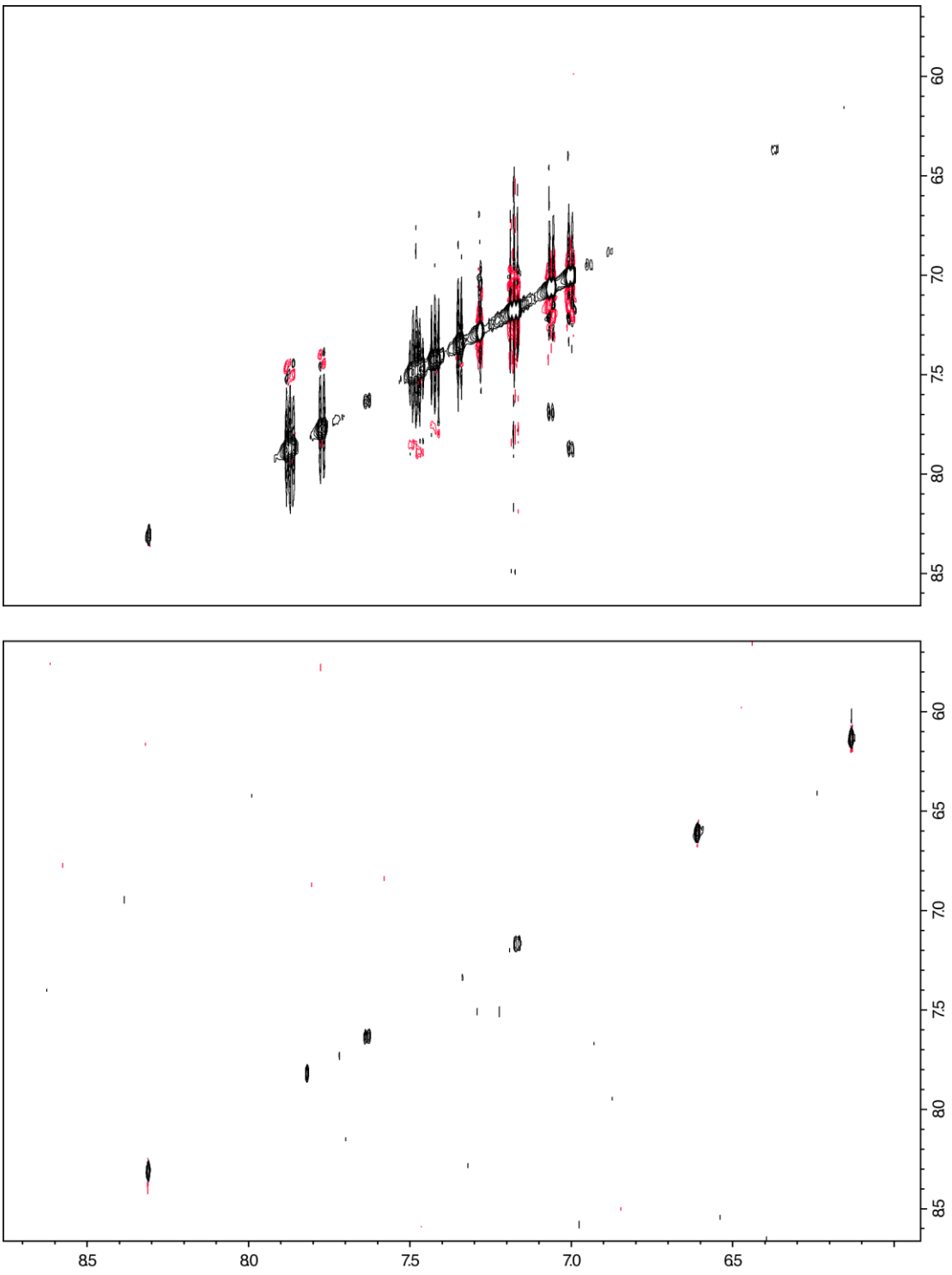


Figure S32: NOESY spectrum of ligand A8 (250 µM) and ligand A4 (750 µM) above and sEH (30 µM) below. These control spectra show that the INPHARMA peaks from Fig. S31 are not protein signals. The mixing time was 800 ms on a 700 MHz spectrometer, equipped with a cryogenically cooled probe head. The spectrum was recorded at 293 K with 64 scans, 8192 points in F2 and 384 points in F1.

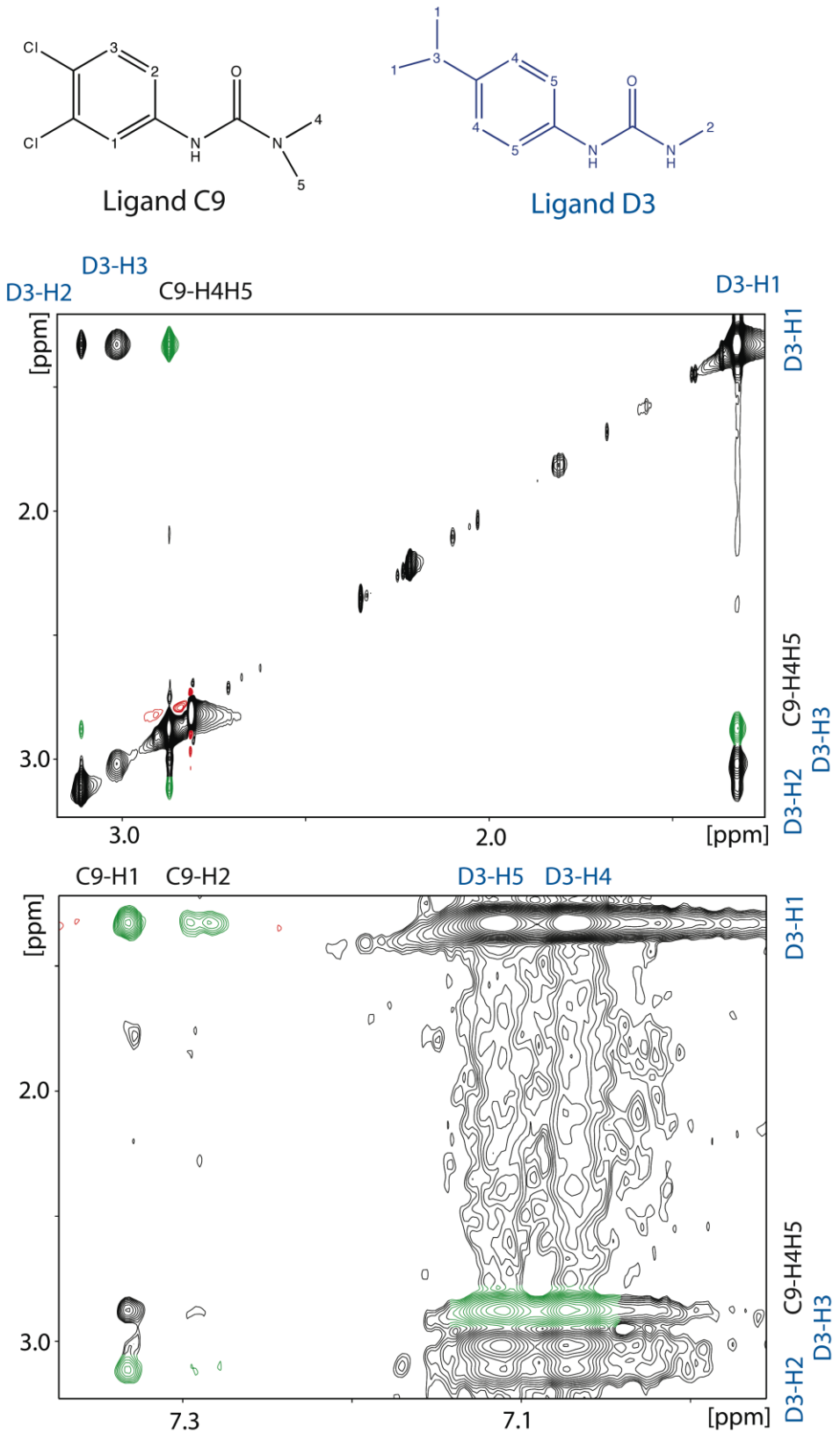
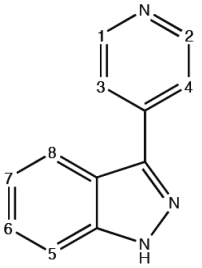


Figure S33: NOESY spectrum of ligand C9 (1 mM) and ligand D3 (1 mM) in the presence of sEH (30 µM). INPHARMA peaks chosen are marked in green. The mixing time was 800 ms on a 700 MHz spectrometer. The spectrum was recorded at 293 K with 128 scans, 4096 points in F2 and 448 points in F1.

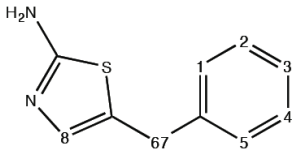
Part IV: Experimental peak volumes

In the following all experimental peak volumes used for the calculations of the three systems PKA, GP and sEH are listed. First, schemes of the ligands of the respective system are given with the naming convention included. Then the STD peak volumes are listed, which were multiplied with 100, as was done within the back-calculation. The trNOE and INPHARMA peak volumes were recorded in one spectrum, therefore they are listed together, with a subdivision as follows: i) first ligand diagonal peak volumes and intra peak volumes, ii) second ligand diagonal peak volumes and intra peak volumes, iii) INPHARMA peak volumes. For the calculations, INPHARMA peaks were normalized on the sum of the diagonal and intra peak volumes in the direct dimension. In the case of caffeine, no diagonal and intra peak volumes were obtained, as the three methyl groups are overlayn with protein methyl signals. The INPHARMA peak volumes for GP were consequently not normalized.

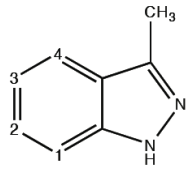
1.PKA ligands



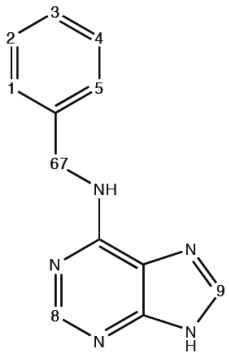
Ligand 1



Ligand 2



Ligand 3



Ligand 4

1.1.PKA STD peak volumes

L1-H1H2	13.75
L1-H8	11.82
L1-H3H4	12.25
L1-H5	10.07
L1-H6	8.91

L2-H2H4	11.12
L2-H1H5	13.67
L2-H8	28.62
L2-H6H7	12.85
L3-H4	9.09
L3-H1	16.92
L3-H2	13.25
L3-CH3	11.84
L4-H8	32.23
L4-H9	32.67
L4-H1H2H3H4H5	5.50

1.2.PKA trNOE and INPHARMA peak volumes of ligand 1 and 2

L1-H1H2	L1-H1H2	66.88
L1-H8	L1-H8	22.46
L1-H3H4	L1-H3H4	61.02
L1-H5	L1-H5	40.99
L1-H6	L1-H6	31.60
L1-H1H2	L1-H8	3.27
L1-H8	L1-H1H2	2.84
L1-H1H2	L1-H3H4	14.50
L1-H3H4	L1-H1H2	13.50
L1-H5	L1-H6	9.33
L1-H6	L1-H5	8.12
L1-H8	L1-H3H4	9.72
L1-H3H4	L1-H8	10.43
L1-H3H4	L1-H5	0.55
L1-H5	L1-H3H4	0.55

L1-H3H4	L1-H6	0.92
L1-H5	L1-H8	0.77
L1-H8	L1-H5	0.81
L1-H8	L1-H6	1.31
L1-H6	L1-H3H4	0.79
L1-H6	L1-H8	1.36
L1-H1H2	L1-H6	0.61
L1-H1H2	L1-H5	0.53
L1-H5	L1-H1H2	0.53
L1-H6	L1-H1H2	0.34
L2-H2H4	L2-H2H4	345.44
L2-H1H3H5	L2-H1H3H5	408.52
L2-H8	L2-H8	142.20
L2-H6H7	L2-H6H7	363.82
L2-H2H4	L2-H1H3H5	28.75
L2-H1H3H5	L2-H2H4	35.37
L2-H2H4	L2-H6H7	3.29
L2-H1H3H5	L2-H6H7	4.02
L2-H8	L2-H6H7	4.43
L2-H6H7	L2-H8	2.47
L1-H1H2	L2-H6H7	0.79
L1-H8	L2-H6H7	0.45
L1-H3H4	L2-H6H7	0.91
L1-H5	L2-H6H7	0.30
L1-H6	L2-H6H7	0.47
L1-H3H4	L2-H1H3H5	0.85
L1-H5	L2-H1H3H5	0.51
L1-H8	L2-H1H3H5	0.59
L1-H1H2	L2-H1H3H5	0.80
L2-H1H3H5	L1-H1H2	0.73

L2-H1H3H5	L1-H8	0.49
L2-H1H3H5	L1-H3H4	0.92
L2-H1H3H5	L1-H5	1.01

1.3. PKA trNOE and INPHARMA peak volumes of ligand 2 and ligand 3

L2-H2H4	L2-H2H4	243.93
L2-H1H3H5	L2-H1H3H5	353.31
L2-H8	L2-H8	109.59
L2-H2H4	L2-H6H7	2.22
L2-H1H3H5	L2-H6H7	8.02
L2-H8	L2-H6H7	8.82
L2-H2H4	L2-H8	0.31
L2-H8	L2-H1H3H5	1.24
L2-H8	L2-H2H4	0.58
L3-H1	L3-H1	142.24
L3-H2	L3-H2	156.11
L3-H3	L3-H3	172.03
L3-H4	L3-H4	175.84
L3-H4	L3-CH3	12.82
L3-H1	L3-CH3	3.15
L3-H2	L3-CH3	1.67
L3-H3	L3-CH3	2.68
L3-H4	L3-H3	17.99
L3-H3	L3-H4	17.45
L3-H2	L3-H3	17.87
L3-H3	L3-H2	17.95
L3-H1	L3-H3	2.76
L3-H3	L3-H1	2.56
L3-H1	L3-H4	1.16
L3-H2	L3-H4	2.27

L3-H1	L3-H2	21.65
L3-H2	L3-H1	20.34
L3-H4	L3-H2	1.97
L3-H4	L3-H1	0.67
L2-H2H4	L3-CH3	1.40
L2-H1H3H5	L3-CH3	2.92
L2-H8	L3-CH3	2.11
L3-H4	L2-H6H7	0.86
L3-H1	L2-H6H7	1.00
L3-H3	L2-H6H7	0.86
L2-H1H3H5	L3-H4	1.06
L3-H2	L2-H6H7	0.58
L3-H4	L2-H1H3H5	1.01
L3-H1	L2-H1H3H5	0.90
L2-H2H4	L3-H4	0.63
L3-H4	L2-H2H4	0.38
L3-H1	L2-H2H4	0.94
L2-H8	L3-H3	0.52
L2-H8	L3-H2	0.43
L2-H8	L3-H4	0.44
L2-H8	L3-H1	0.53

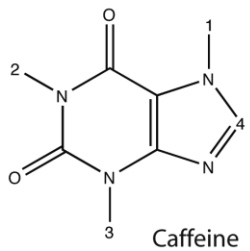
1.4. PKA trNOE and INPHARMA peak volumes of ligand 3 and ligand 4

L3-H4	L3-H4	502.24
L3-H1	L3-H1	440.12
L3-H2	L3-H2	503.69
L3-H3	L3-H3	526.48
L3-H4	L3-CH3	10.77
L3-H2	L3-CH3	3.17
L3-H1	L3-CH3	3.68

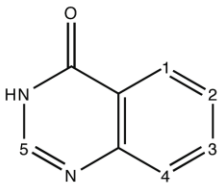
L3-H3	L3-CH3	4.75
L3-H2	L3-H3	12.77
L3-H3	L3-H2	13.85
L3-H4	L3-H3	11.91
L3-H3	L3-H4	11.47
L3-H1	L3-H3	3.01
L3-H4	L3-H1	0.85
L3-H3	L3-H1	2.92
L4-H8	L4-H8	280.08
L4-H9	L4-H9	268.29
L4-H1H5	L4-H1H5	997.21
L4-H2H4	L4-H2H4	884.89
L4-H3	L4-H3	402.93
L4-H8	L4-H1H5	4.58
L4-H9	L4-H1H5	2.95
L4-H1H5	L4-H8	2.29
L4-H2H4	L4-H8	2.09
L4-H8	L4-H2H4	1.26
L4-H9	L4-H2H4	1.36
L4-H8	L3-CH3	6.82
L4-H9	L3-CH3	2.54
L4-H1H5	L3-CH3	4.86
L4-H2H4	L3-CH3	4.82
L4-H3	L3-CH3	2.86
L4-H8	L3-H3	1.04
L4-H9	L3-H3	1.12
L4-H8	L3-H1	1.07
L4-H9	L3-H1	1.88
L4-H8	L3-H2	1.31
L4-H9	L3-H2	1.24

L4-H8	L3-H4	0.83
L4-H9	L3-H4	0.52
L3-H4	L4-H1H5	2.80
L4-H1H5	L3-H4	2.22
L4-H2H4	L3-H4	2.86
L4-H3	L3-H4	0.81
L3-H4	L4-H8	0.65
L3-H1	L3-H4	1.62
L3-H3	L4-H9	0.72
L3-H1	L4-H9	1.78
L3-H3	L4-H8	0.79

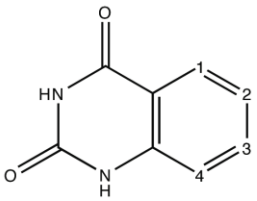
2.GP ligands



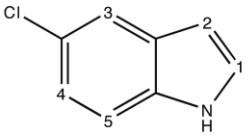
Caffeine



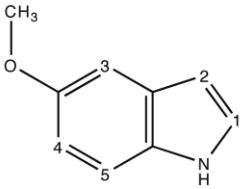
Ligand 5



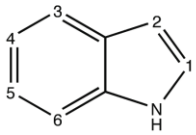
Ligand 7



5-chloroindole



Ligand 6



Ligand 8

2.1.GP STD peak volumes

Caffeine-H4	3.24
Caffeine-H1	8.94
Caffeine-H3	19.34
Caffeine-H2	12.99

5-chloroindole-H3	69.50
5-chloroindole-H5	64.20
5-chloroindole-H1	53.35
5-chloroindole-H4	68.17
5-chloroindole-H2	66.83
5-H1H5	25.54
5-H3	37.30
5-H4	40.08
5-H2	21.17

2.2. GP trNOE and INPHARMA peak volumes of caffeine and ligand 5

5-H1H5	5-H1H5	466.3743896
5-H3	5-H3	223.4978333
5-H4	5-H4	220.3064117
5-H2	5-H2	219.1618958
5-H1H5	5-H2	12.18059349
5-H2	5-H1H5	13.91180134
5-H3	5-H2	10.54667854
5-H2	5-H3	13.00882053
5-H4	5-H2	2.257376909
5-H2	5-H4	6.160551548
5-H3	5-H4	32.78610229
5-H4	5-H3	26.58098793
5-H1H5	5-H3	1.322131038
5-H3	5-H1H5	2.061654091
5-H1H5	5-H4	0.927576959
5-H4	5-H1H5	0.654471457
5-H4	Caffeine-H3	0.333972782

5-H4	Caffeine-H2	0.286721617
5-H4	Caffeine-H1	0.249226242
Caffeine-H1	5-H4	0.424235195
Caffeine-H2	5-H4	0.321579605
Caffeine-H3	5-H4	0.506145835
5-H3	Caffeine-H3	0.472887158
5-H3	Caffeine-H2	0.301943898
5-H2	Caffeine-H3	0.190084234
5-H2	Caffeine-H2	0.143963575

2.3. GP trNOE and INPHARMA peak volumes of caffeine and ligand 7

7-H1	7-H1	138.6291656
7-H3	7-H3	141.9783173
7-H2	7-H2	144.2865143
7-H4	7-H4	158.1337585
7-H3	7-H2	6.512661457
7-H2	7-H3	6.696210384
7-H3	7-H4	8.526493073
7-H4	7-H3	8.67834568
7-H1	7-H2	4.580965519
7-H2	7-H1	5.103363514
7-H1	7-H3	0.672525942
7-H3	7-H1	0.5304721
Caffeine-H1	7-H1	0.546867192
Caffeine-H2	7-H1	0.653225482
Caffeine-H3	7-H1	0.577695131

2.4. GP trNOE and INPHARMA peak volumes of 5-chloroindole and ligand 6

5-chloroindole-H3	5-chloroindole-H3	15.37984657
5-chloroindole-H5	5-chloroindole-H5	15.92882538
5-chloroindole-H1	5-chloroindole-H1	23.40568352
5-chloroindole-H4	5-chloroindole-H4	12.91070938
5-chloroindole-H2	5-chloroindole-H2	87.56942749
5-chloroindole-H5	5-chloroindole-H4	5.735244274
5-chloroindole-H4	5-chloroindole-H5	5.351717949
5-chloroindole-H3	5-chloroindole-H2	2.945459366
5-chloroindole-H2	5-chloroindole-H3	3.248150587
5-chloroindole-H1	5-chloroindole-H2	3.727215767
5-chloroindole-H2	5-chloroindole-H1	4.892887115
6-H5	6-H5	122.8669281
6-H1	6-H1	121.5041656
6-H3	6-H3	127.2133255
6-H4	6-H4	117.533699
6-H2	6-H2	123.1285782
6-CH3	6-CH3	887.62854
6-H5	6-H4	12.56416893
6-H4	6-H5	12.4762373
6-H3	6-H2	2.75953269
6-H2	6-H3	3.454149246
6-H1	6-H2	6.952414989
6-H2	6-H1	6.500362396
6-H2	6-CH3	2.283021927
6-H4	6-CH3	2.499403238
6-H3	6-CH3	10.04872799
6-H5	6-CH3	2.390412331
6-CH3	6-H2	2.789409399
6-CH3	6-H4	3.735375404
6-CH3	6-H3	13.50357151

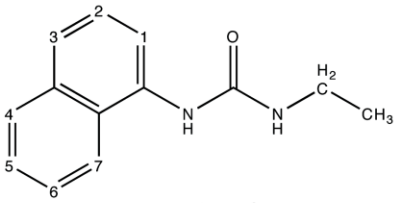
6-CH3	6-H1	1.167292714
6-CH3	6-H5	2.190643311
5-chloroindole-H3	6-CH3	0.95917511
5-chloroindole-H4	6-CH3	0.303892225
5-chloroindole-H3	6-H3	0.849735856
6-H3	5-chloroindole-H3	0.602896035
5-chloroindole-H4	6-H4	0.228223145
6-H4	5-chloroindole-H4	0.204516605
6-H4	5-chloroindole-H3	0.371615887

2.5. GP trNOE and INPHARMA peak volumes of 5-chloroindole and ligand 8

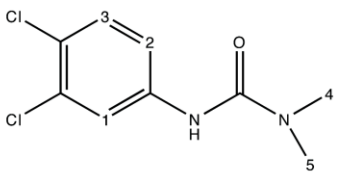
5-chloroindole-H5	5-chloroindole-H5	10.90097904
5-chloroindole-H4	5-chloroindole-H4	10.3770008
5-chloroindole-H2	5-chloroindole-H2	17.27019691
5-chloroindole-H5	5-chloroindole-H4	7.53403759
5-chloroindole-H4	5-chloroindole-H5	5.131019115
8-H6	8-H6	170.4409027
8-H5	8-H5	162.317749
8-H4	8-H4	147.5278473
8-H2	8-H2	177.5836945
8-H6	8-H5	14.20987988
8-H5	8-H6	14.82373238
8-H6	8-H4	2.91305232
8-H4	8-H6	2.915114164
8-H5	8-H4	30.80410004
8-H4	8-H5	36.69189835
8-H5	8-H2	1.596555948
8-H2	8-H5	0.928431749

8-H4	8-H2	1.609217048
8-H2	8-H4	1.726322174
8-H6	8-H2	1.222156525
8-H2	8-H6	1.075672746
5-chloroindole-H5	8-H6	1.806460977
8-H6	5-chloroindole-H2	0.366241276
8-H4	5-chloroindole-H2	0.552625597

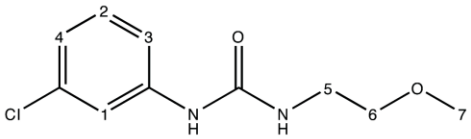
3. sEH ligands



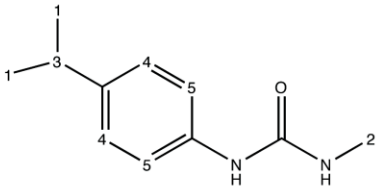
Ligand A8



Ligand C9



Ligand A4



Ligand D3

3.1. sEH STD peak volumes

A8-H1H2H5H6	38.25
A8-H3	38.85
A8-H4H7	29.07
A8-CH2	58.38
A8-CH3	42.18
D3-H5	15.85

D3-H4	15.40
D3-H2	13.90
D3-H1	12.74
D3-H3	16.30
A4-H1	34.16
A4-H2	50.96
A4-H3	27.87
A4-H4	25.05
A4-H6	26.71
A4-H5H7	50.41
C9-H1H2	31.55
C9-H3	30.10
C9-H4H5	30.57

3.2. sEH trNOE and INPHARMA peak volumes of ligand A8 and ligand A4

A8-H7	A8-H7	14.04
A8-H4	A8-H4	6.33
A8-H3	A8-H3	6.89
A8-H5H6	A8-H5H6	29.95
A8-H2	A8-H2	12.8
A8-H1	A8-H1	16.95
A8-CH2	A8-CH2	39.25
A8-CH3	A8-CH3	71.53
A8-H7	A8-H5H6	12.22
A8-H5H6	A8-H7	13.11
A8-CH3	A8-CH2	23.27
A8-CH2	A8-CH3	23.74
A4-H1	A4-H1	1031.29

A4-H2	A4-H2	650.08
A4-H3	A4-H3	894.72
A4-H4	A4-H4	766.14
A4-H6	A4-H6	2174.8
A4-H5H7	A4-H5H7	6998.35
A4-H6	A4-H5H7	618.52
A4-H5H7	A4-H6	608.8
A4-H1	A4-H2	13.32
A4-H2	A4-H1	17.42
A4-H1	A4-H3	25.81
A4-H3	A4-H1	21.05
A4-H1	A4-H4	21.12
A4-H4	A4-H1	20.13
A4-H2	A4-H3	252.73
A4-H3	A4-H2	162.61
A4-H2	A4-H4	163.12
A4-H4	A4-H2	198.36
A4-H3	A4-H4	27.69
A4-H4	A4-H3	39.92
A4-H1	A4-H5H7	108.37
A4-H2	A4-H5H7	91.36
A4-H3	A4-H5H7	90.91
A4-H4	A4-H5H7	89.36
A4-H1	A4-H6	43.7
A4-H2	A4-H6	38.4
A4-H3	A4-H6	39.36
A4-H4	A4-H6	35.2
A4-H6	A8-CH3	6.18
A8-CH3	A4-H4	12.65
A8-CH3	A4-H5H7	35.74
A4-H1	A8-CH3	11.17

A4-H2	A8-CH3	7.05
A4-H3	A8-CH3	7.43
A4-H4	A8-CH3	10.1
A8-H7	A4-H5H7	15.25
A8-H4	A4-H5H7	5.77
A8-H3	A4-H5H7	8.01
A8-H5H6	A4-H5H7	26.69
A8-H2	A4-H5H7	19.1
A8-H1	A4-H5H7	15.91
A8-H7	A4-H6	6.17
A8-H4	A4-H6	1.6
A8-H3	A4-H6	3.84
A8-H5H6	A4-H6	11.05
A8-H2	A4-H6	5.91
A8-H1	A4-H6	6.49
A8-H7	A4-H2	2.91
A8-H5H6	A4-H3	3.75
A4-H3	A8-H5H6	5.25
A8-H5H6	A4-H2	4.41
A4-H2	A8-H5H6	2.9

3.3. sEH trNOE and INPHARMA peak volumes of ligand A4 and ligand C9

A4-H1	A4-H1	1325.35
A4-H2	A4-H2	736.49
A4-H3	A4-H3	1079.29
A4-H4	A4-H4	943.09
A4-H6	A4-H6	2662.34
A4-H5H7	A4-H5H7	9288.75
A4-H1	A4-H5H7	161.27
A4-H2	A4-H5H7	133.56

A4-H3	A4-H5H7	156.04
A4-H4	A4-H5H7	128.52
A4-H1	A4-H6	69.16
A4-H2	A4-H6	52.73
A4-H3	A4-H6	63.06
A4-H4	A4-H6	48
A4-H6	A4-H1	44.22
A4-H5H7	A4-H1	141.87
A4-H6	A4-H2	36.05
A4-H5H7	A4-H2	125.65
A4-H6	A4-H3	39.95
A4-H5H7	A4-H3	133.14
A4-H6	A4-H4	28.95
A4-H5H7	A4-H4	96.71
A4-H6	A4-H5H7	1143.37
A4-H5H7	A4-H6	1242.98
A4-H2	A4-H3	413.69
A4-H3	A4-H2	425.33
A4-H2	A4-H4	287.68
A4-H4	A4-H2	322.79
A4-H1	A4-H4	51.87
A4-H4	A4-H1	63.9
A4-H1	A4-H2	49.74
A4-H2	A4-H1	65.78
A4-H1	A4-H3	67.54
A4-H3	A4-H1	102.32
C9-H1H2	C9-H1H2	205.04
C9-H4H5	C9-H4H5	1340.06
C9-H1H2	C9-H4H5	27.8
A4-H1	C9-H4H5	36.62

A4-H2	C9-H4H5	32.94
A4-H3	C9-H4H5	42.95
A4-H4	C9-H4H5	31.27
C9-H1H2	A4-H5H7	61.92
C9-H1H2	A4-H6	28.36
C9-H1H2	C9-H3	68.26
A4-H6	C9-H4H5	32.73
C9-H4H5	A4-H6	42.51
A4-H5H7	C9-H4H5	112.28
C9-H4H5	A4-H5H7	125.01
C9-H1H2	A4-H2	16.55

3.4. sEH trNOE and INPHARMA peak volumes of ligand C9 and ligand D3

C9-H4H5	C9-H4H5	1865.12
C9-H1	C9-H1	144.77
C9-H2	C9-H2	85.46
C9-H2	C9-H3	26.09
C9-H3	C9-H2	20.69
D3-H1	D3-H1	11708.78
D3-H2	D3-H2	5749.53
D3-H3	D3-H3	928.08
D3-H4	D3-H4	4002.21
D3-H5	D3-H5	4696.6
D3-H2	D3-H1	1090.54
D3-H1	D3-H2	1028.78
D3-H3	D3-H1	1501.63
D3-H1	D3-H3	1650.83
D3-H4	D3-H1	1368.2
D3-H5	D3-H1	1437.26

D3-H4	D3-H2	371.06
D3-H5	D3-H2	441.91
D3-H4	D3-H3	500.36
D3-H5	D3-H3	461.52
D3-H2	D3-H4H5	878.95
D3-H3	D3-H4H5	840.82
D3-H1	D3-H4H5	2284.76
C9-H4H5	D3-H1	506.67
D3-H1	C9-H4H5	481.02
C9-H4H5	D3-H2	129.52
D3-H2	C9-H4H5	345.01
C9-H4H5	D3-H3	126.75
C9-H1	D3-H1	64.47
C9-H1	D3-H2	36.37
C9-H1	C9-H4H5	42.48
C9-H2	D3-H1	56.46
C9-H2	D3-H2	21.69
C9-H2	C9-H4H5	23.73
D3-H4	C9-H4H5	206.9
D3-H5	C9-H4H5	225.55
C9-H4H5	D3-H4H5	463.23
C9-H4H5	C9-H1H2	63.14
D3-H2	C9-H1H2	97.82
D3-H1	C9-H1H2	313.3
C9-H1	D3-H4H5	34.03

References

[1] C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Cryst.* **2011**, D67, 293-302;

[2] W. J. Kabsch, *Appl. Cryst.* **1993**, 26, 795-800;

[3] G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi , A. Sharff, O. S. Smart, C. Vonrhein, T. O. Womack, **2011**, BUSTER version 2.9.4 Cambridge, United Kingdom: Global Phasing Ltd.;

[4] P. Emsley, K. Cowtan, *Acta Cryst*. **2004**, D60, 2126-2132;

[5] G. A. Gomez, C. Morisseau, B. D. Hammock, D. W. Christianson,.*Biochemistry*, **2004**;43, 4716-4723;

[6] A. Cronin, S. Mowbray, H. Dürk, S. Homburg, I. Fleming, B. Fisslthaler, F. Oesch, M. Arand, *Proc. Natl. Acad. Sci. USA.* **2003**, 100**,** 1552-1557;

[7] F. W. Studier, *Prot. Expr. Purif.* **2005**, 41**,** 207-234;

[8] P.R.Evans, *Proceedings of CCP4 Study Weekend, 1993, on Data Collection Processing*, **1993**, 114-122;

[9] Collaborative Computational Project, Number 4. *Acta Cryst*. **1994**, D50, 760-763;

[10] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Cryst.* **2007**, 40, 658-674;