

Crystallography-Independent Determination of Ligand Binding Modes**

Julien Orts, Jennifer Tuma, Marcel Reese, S. Kaspar Grimm, Peter Monecke, Stefan Bartoschek, Alexander Schiffer, K. Ulrich Wendt,* Christian Griesinger,* and Teresa Carlomagno*

Within the last few decades, structure-based drug design (SBDD) has evolved into a powerful tool for the optimization of many low-molecular-weight lead compounds into highly potent drugs.^[1] The principle of SBDD lies in the combination of different chemical moieties with the aim of obtaining a molecule that, while possessing the pharmacological properties necessary for a drug, is complementary in shape to the receptor binding pocket. This process requires knowledge of the exact structure of the protein/ligand complex. At present, structural genomics initiatives provide protein structures of biomedically relevant targets at an increasing rate^[2] and recent structures of ion channels^[3] and G-protein-coupled receptors (GPCRs)^[4–6] bring these protein classes within reach for SBDD. Despite these successes, the daily work of pharmaceutical discovery is often limited by the ability to obtain high-resolution crystal structures of the target proteins in complex with the lower affinity ligands (lead structures) that are commonly identified by high-throughput screening or

by fragment-based lead discovery.^[1] In view of this limitation, SBDD would benefit from methods providing the relative orientations of different chemical fragments binding competitively to a receptor site. Such an approach would provide protein/ligand structures of novel ligands or fragments in relation to the known cocrystal structure of a reference ligand.

Recently, we reported the observation by NMR spectroscopy of interligand NOE peaks occurring between two small ligands binding weakly and competitively to the same binding pocket of a common macromolecular receptor (Figure 1).^[7,8]

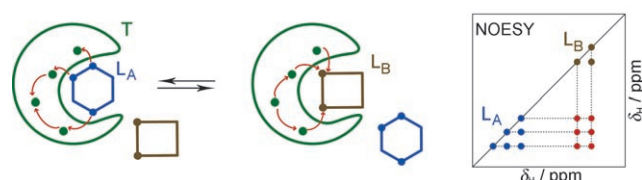


Figure 1. Principle of the INPHARMA signals. The mechanisms of magnetization transfer between the two ligands, L_A and L_B , can be described as follows: During the mixing time of a NOESY experiment, L_A binds to the receptor and its proton H_A transfers its magnetization to receptor proton H_T . Subsequently, L_A dissociates from the receptor and L_B binds. The magnetization previously transferred from H_A to H_T can now be transferred from H_T to H_B , a proton of L_B . This leads to an intermolecular NOE cross-peak between H_A and H_B that has been mediated by the receptor proton H_T (spin diffusion).

The measured mixture in solution contained two ligands (L_A and L_B) in a 10- to 50-fold excess relative to the target receptor (T). As the ligands were competitive binders, these NOEs did not originate from a direct transfer of magnetization between the two ligands,^[9] but rather from a spin-diffusion process mediated by the protons of the receptor binding pocket. We proposed that such interligand NOEs can be used to define the relative orientation of the two ligands in the receptor binding pocket (the relative binding mode) and we termed the novel effect INPHARMA (internuclear NOEs for pharmacophore mapping;^[8] Figure 1). INPHARMA can be observed for complexes with a dissociation constant (K_d) in the low micromolar to millimolar range.

The scope of this work is to demonstrate for the first time that the INPHARMA method allows the determination of the relative, and in favorable cases even the absolute, binding mode of two low-affinity ligands binding competitively to a common receptor site and that it can thus be applied in the context of SBDD. In accordance with existing SBDD workflows, the experimental information derived from INPHARMA was used to select the correct binding mode

[*] Dr. P. Monecke, Dr. S. Bartoschek, Dr. A. Schiffer, Dr. K. U. Wendt
Department of Chemical and Analytical Sciences/Structural Biology
Sanofi-Aventis Deutschland GmbH
R&D CAS Structural Biology FFM
Industriepark Hoechst, Bldg. G877
65926 Frankfurt am Main (Germany)
E-mail: ulrich.wendt@sanofi-aventis.com

J. Orts,^[†] Dr. J. Tuma,^[†] M. Reese, Dr. S. K. Grimm,
Prof. Dr. C. Griesinger
Department of NMR-based Structural Biology
Max Planck Institute for Biophysical Chemistry
Am Fassberg, 37077 Göttingen (Germany)
E-mail: cigr@nmr.mpibpc.mpg.de

J. Orts,^[†] Dr. T. Carlomagno
Structural and Computational Biology Unit
European Molecular Biology Laboratory (EMBL)
Meyerhofstrasse, 69117 Heidelberg (Germany)
Fax: (+49) 6221-387-8519
E-mail: teresa.carlomagno@embl.de

[†] These authors contributed equally to this work.

[**] This work was supported by the Volkswagen-Stiftung (grant to T.C.), the Fonds der chemischen Industrie (fellowship no. 180081 to J.O.; support to C.G.), the Max-Planck Gesellschaft, and Sanofi-Aventis. We thank Elke Duchardt for the measurement of some interligand NOE spectra, Sandra Schimanski-Breves for crystallization, Thomas Langer for protein production of Chinese hamster protein kinase A (PKA), and Prof. H. Schwalbe for the provision of previously reported expression plasmids. Bovine PKA was obtained from Dirk Bossemeyer and Michael Gassel. Crystallographic data were collected on beamline ID14.1 at the European Synchrotron Radiation Facility (ESRF).

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.200801792>.

the experimental and theoretical INPHARMA peaks but that show systematic deviations (under- or overestimation) of groups of INPHARMA peaks stemming from separate structural parts of the ligand(s).

- 3) Semiquantitative use of weak INPHARMA peaks that are observable either in longer experiments at higher magnetic fields or at higher protein concentrations and/or with longer mixing times. At this step, weak additional INPHARMA peaks are used as a discrimination criterion. However, the quantitative fitting of weak INPHARMA signals is often deteriorated by the effect of internal motions in multiple steps of spin diffusion; thus, these NOEs are interpreted in a semiquantitative manner.

The correlation coefficients of the calculated versus the measured INPHARMA signals for all 16 pairs of possible binding orientations are given in Table S2 in the Supporting Information. We found that the correlation coefficients for 4 out of the 16 pairs are clearly different from the average correlation and, therefore, we selected these 4 model pairs with $R^2 > 0.9$ or $Q^2 < 0.15$ to pass the first selection criterion (Figure 4). In all of the model pairs A–D, the orientation of L_B is uniquely defined (Figure 4), whereas the orientation of L_A is largely undefined. This finding most likely results from the different three-dimensional shape of the two ligands. For flat

L_A , the intermolecular proton–proton distance distribution does not substantially change upon rotations of the ligand in the binding pocket (Figure S1 in the Supporting Information); this translates into a reduced discriminatory power of the INPHARMA method towards different orientations of L_A .

To further differentiate amongst the four ligand orientations that passed the first selection, we analyzed the fit of the calculated and experimental INPHARMA signals of individual ring systems (criterion 2). As seen in Figure 4, pairs A and B perform significantly better than pairs C and D (Table S2 in the Supporting Information). For these last two pairs, the INPHARMA signals stemming from the phenyl ring of L_B to the pyridine ring of L_A (Figure 4, yellow; Figure S2 in the Supporting Information) and to the 3-(4-pyridyl)indazole ring of L_A (Figure 4, red; Figure S2 in the Supporting Information) are consistently over- and underestimated, respectively. These results indicate that, in pairs C and D, the orientation of L_A is wrong by a 180° rotation around the y axis. Thus, model pairs C and D can be safely excluded.

The degeneracy of the H1/2 and H3/4 protons of L_A and the H1–5 protons of L_B impedes the discrimination by criteria 1 and 2 between the remaining pairs, A and B, which differ by a rotation of L_A around the z axis. In accordance with selection criterion 3, we recorded an additional NOESY spectrum on a 900 MHz spectrometer, with double the number of scans and a mixing time of 600 ms, with the aim of collecting more experimental data from the nondegenerated protons (Figure S3 in the Supporting Information). This spectrum indeed showed additional interligand NOEs, between protons H6 and H7 of L_B and protons H5, H6, and H8 of L_A . The ratio of 1.4 between the H8(L_A)–H6/7(L_B) NOE and the H5(L_A)–H6/7(L_B) NOE is well reproduced in the back-calculated interligand NOEs for model pair A (1.8), whereas it is dramatically underestimated in model pair B (0.06); thus, the latter pair can be excluded.

Application of selection criteria 1–3 allowed us to identify model pair A as the one uniquely representing the experimental INPHARMA data. A comparison of model pair A with the crystal structures of the PKA/ L_A and PKA/ L_B complexes shows that this model reproduces both the correct relative orientation and the correct absolute orientation of the two ligands with respect to the protein. The main scope of the methodology is to derive the relative binding mode of two low-affinity fragment inhibitors; in the test case shown here, the INPHARMA approach exceeds the expectations and allows the determination of the absolute binding mode of both ligands. In general, we expect this to occur when the structure of the apoprotein closely resembles that of the protein in the complex.

Further docking models, described in Figure S4 in the Supporting Information and providing a more complete sampling around some of the orientations illustrated in Figure 3, were tested as well, which led to the selection of the same crystal-structure-like relative orientation of the two ligands in the PKA binding pocket as that in model pair A. Finally, we verified that the INPHARMA peaks calculated for the crystal structures of the PKA/ L_A and PKA/ L_B complexes reproduce the experimental INPHARMA peaks

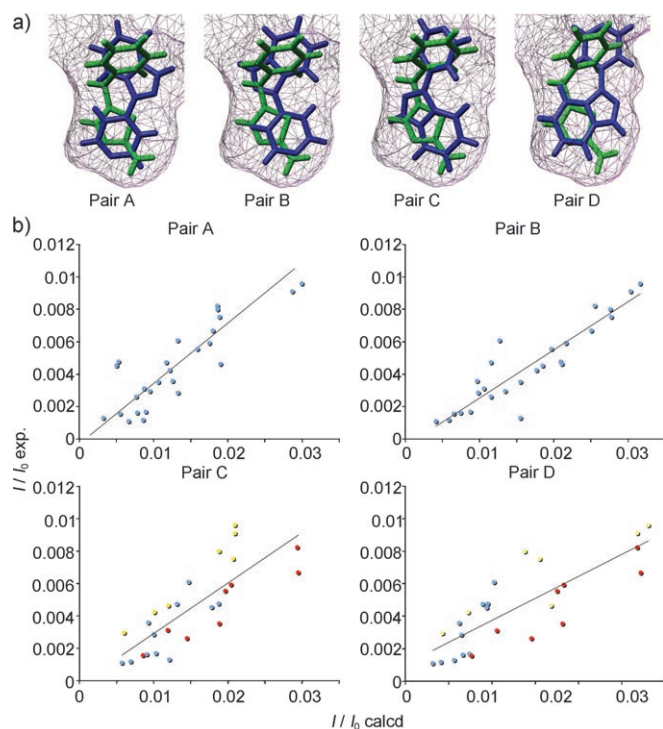


Figure 4. a) Representation of the four model pairs (A–D) that show good correlation between the theoretical and the experimental INPHARMA signals. b) Correlations of the normalized experimental INPHARMA peaks with the back-calculated INPHARMA peaks, for each model pair from (a). The INPHARMA signals from the phenyl ring of L_B to either the pyridine or the 3-(4-pyridyl)indazole ring of L_A for models C and D (color-coded in yellow and red, respectively; see also Figure S2 in the Supporting Information) are consistently under- or overestimated for the model pairs C and D.

in a comparable manner to those for model pair A (Figure S5 in the Supporting Information). The fitting of the calculated versus experimental INPHARMA signals was excellent. The residual deviations can be accounted for by invoking the presence of internal motions,^[18,19] as proven by the slope of the linear regression of 0.3 (see Figure S6 in the Supporting Information).^[20,21] The good quality of the fit further verifies that the INPHARMA peaks are suitable experimental data for ranking of the docking poses.

In the case of PKA, the structure of the protein binding pocket in the various PKA/ligand complexes is largely unchanged and the PKA conformation used to calculate the different docking models is quite close to the “true” structure of the protein in the complexes (root mean square deviation (rmsd) of the binding pocket heavy atoms ≈ 0.3 Å). However, this might not be the case for other systems, in which the binding pocket might be dynamic or ill defined in the available structural models. In order to assess the influence of the quality of the protein structure on the outcome of the methodology, we used molecular dynamics simulations to generate a series of PKA structures with inaccurate conformations of the binding pocket. The models exhibit heavy-atom rmsd values of 1.9, 1.7, 1.9, 1.6, and 1.9 Å with respect to the X-ray structure of the binding pocket. Ligands L_A and L_B were docked in the 4 different orientations described above and sets of 16 model pairs were generated for each of the 5 PKA structures (Figure S7 in the Supporting Information). The calculated INPHARMA peaks for each of the 16×5 model pairs were correlated with the experimental interligand NOEs as described above (Table S3 in the Supporting Information). We found that the correct absolute orientation of the two ligands was selected by the criteria described above for four of the five inaccurate protein models, whereas pose 4 (or pose 2; Figure 3b) was selected for both ligands in the fifth PKA model. Still, even for this model, the correct relative orientation of L_A and L_B is predicted. This test showed that prediction of the relative binding mode of the two ligands is robust towards inaccuracy in the applied protein structure; thus the INPHARMA method may allow the alignment of chemically different ligands as a basis for pharmacophore modeling in the absence of a protein crystal structure. On the other hand, prediction of the correct absolute orientation of the ligands depends on the accuracy of the protein structure model used during docking.

The INPHARMA method closes a gap in structure-based drug discovery, in which, commonly, crystal structures of some but not all interesting chemical lead series are accessible. We have demonstrated that the INPHARMA method is a solid experimental approach to determine the complex structure of novel interesting lead series relative to a known protein/ligand crystal structure. The bound ligand structure, easily obtainable from transferred NOE data,^[22,23] and a structural model of the apoprotein, which together are used to generate plausible docking modes, are the only requirements for the application of the method. The possibility of determining absolute protein/ligand orientations in the absence of a crystal structure of the complex suggests novel approaches for structure-based drug design, in which INPHARMA data for an appropriate selection of differently shaped low-molecular-

weight ligands may provide a surprisingly detailed ligand-based map of the receptor binding site.

Experimental Section

Protein expression and purification for the NMR experiments: The Chinese hamster α catalytic subunit of cyclic adenosine monophosphate (cAMP) dependent protein kinase A (PKA) was expressed and purified according to the published procedure.^[24]

NMR experiments: NOESY experiments (20 h each) were performed with a mixture of L_A , L_B , and PKA on an 800 MHz spectrometer by using a cryogenically cooled probe head ($\tau_m = 150$, 300, 450, 600, and 750 ms; Table S1 in the Supporting Information). Data were recorded for two different samples: for $\tau_m = 300$ and 600 ms, $[L_A] = 150$ μ M, $[L_B] = 450$ μ M, and $[PKA] = 30$ μ M; for $\tau_m = 150$, 450, and 750 ms, $[L_A] = 150$ μ M, $[L_B] = 450$ μ M, and $[PKA] = 25$ μ M. The different protein concentrations in the two samples were accounted for in the calculation of the theoretical INPHARMA peaks. An additional NOESY spectrum was acquired on a 900 MHz spectrometer with $\tau_m = 600$ ms. The sample contained 150 μ M $[L_A]$, 450 μ M $[L_B]$, and 25 μ M $[PKA]$.

Crystal structures: The recombinant bovine α catalytic subunit of cAMP-dependent protein kinase was crystallized according to Engh et al.^[14] Inhibitor complexes were obtained by soaking the crystals for 24–48 h in the crystallization solution with 10 mM L_A or L_B . Diffraction data were collected on beamline ID14.1 at ESRF. Datasets were processed, the structure was solved and initially refined by using APRV,^[25] XDS,^[26] and CNX software (Accelrys, San Diego). 7% of the reflections were set aside for cross-validation. Interactive model building and refinement was carried out by using O,^[27] COOT,^[28] CNX, and REFMAC5^[29] software to yield final (R/R_{free}) values (R is the R factor $(\sum |F_o| - \sum |F_c|)/\sum |F_o|$; R_{free} is the R factor calculated using 5% of the data that were excluded from the refinement) of 21.5/24.9% for the PKA: L_A and 22.0/26.3% for the PKA: L_B complex (Table S4 in the Supporting Information). Coordinates and structure factors are available at the Protein Data Bank (PDB) under entry codes 3DNE.pdb (L_A) and 3DND.pdb (L_B).

Docking models: Alternative binding poses were generated as follows. By starting with the crystal structure, the ligands were rotated by 180° either around the y or the z axis, or both (Figure 3), and the structures were subsequently energy minimized with Xplor software,^[30] with the ligand kept rigid and the repulsive part of the Van der Waals potential used as the driving force. A rotation of 90° around the y or z axis was prohibited by the shape of the binding pocket.

Calculation of the INPHARMA peaks: The INPHARMA signals were calculated for all pairs of protein/ligand structures resulting from the obtained relative orientations, by using the full-relaxation matrix approach, as described in the Supporting Information.^[7] All protons at a distance of $d < 1$ nm from any ligand proton were included in the calculations.^[7] The relative affinity of the two ligands was obtained by competition against Y-27632 (ROCK inhibitor), which delivered an inhibition constant (K_i) of 6 μ M for L_A and 16 μ M for L_B . Therefore, a ratio of 1:3 was used for k_{AB}/k_{BA} [Eq. (S3) in the Supporting Information]. The INPHARMA peaks were normalized by the diagonal peak with an equal (ω_2) frequency in the direct dimension at the lowest mixing time. The correlation time of the complex was optimized by fitting the experimental intraligand transferred NOEs to their theoretical values. Internal motions of the ligand in the binding pocket or of the protein side chains are ignored in the calculation, due to the lack of a suitable theoretical model. The presence of internal motion results in a slope of the linear regression between the experimental and calculated INPHARMA signals that is different from 1 at shorter mixing times. At very long mixing times, where several different pathways of magnetization transfer come into play, the presence of internal motions of variable amplitude at different sites in the complex is expected to deteriorate the overall quality of the fit.

MD simulations: A molecular dynamics protocol at 298 K for 15 ns in a continuum matter (dielectric constant = 78.4) was used to generate PKA conformations other than that determined by the X-ray crystal structure. One structure was saved every 3 ns for a total of 5 different structures (Figure S7 in the Supporting Information).

Received: April 17, 2008

Published online: September 2, 2008

Keywords: drug design · kinases · ligand binding · NMR spectroscopy · pharmacophore mapping

- [1] D. C. Rees, M. Congreve, C. W. Murray, R. Carr, *Nat. Rev. Drug Discovery* **2004**, *3*, 660.
- [2] B. W. Matthews, *Nat. Struct. Mol. Biol.* **2007**, *14*, 459.
- [3] F. M. Ashcroft, *Nature* **2006**, *440*, 440.
- [4] D. M. Rosenbaum, V. Cherezov, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, X. J. Yao, W. I. Weis, R. C. Stevens, B. K. Kobilka, *Science* **2007**, *318*, 1266.
- [5] V. Cherezov, D. M. Rosenbaum, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka, R. C. Stevens, *Science* **2007**, *318*, 1258.
- [6] S. G. Rasmussen, H. J. Choi, D. M. Rosenbaum, T. S. Kobilka, F. S. Thian, P. C. Edwards, M. Burghammer, V. R. Ratnala, R. Sanishvili, R. F. Fischetti, G. F. Schertler, W. I. Weis, B. K. Kobilka, *Nature* **2007**, *450*, 383.
- [7] M. Reese, V. M. Sánchez-Pedregal, K. Kubicek, J. Meiler, M. J. J. Blommers, C. Griesinger, T. Carlomagno, *Angew. Chem.* **2007**, *119*, 1896; *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 1864.
- [8] V. M. Sánchez-Pedregal, M. Reese, J. Meiler, M. J. Blommers, C. Griesinger, T. Carlomagno, *Angew. Chem.* **2005**, *117*, 4244; *Angew. Chem. Int. Ed.* **2005**, *44*, 4172.
- [9] R. E. London, *J. Magn. Reson.* **1999**, *141*, 301.
- [10] M. J. Stocks, S. Barber, R. Ford, F. Leroux, S. St-Gallay, S. Teague, Y. Xue, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3459.
- [11] D. R. Knighton, J. H. Zheng, L. F. Ten Eyck, V. A. Ashford, N. H. Xuong, S. S. Taylor, J. M. Sowadski, *Science* **1991**, *253*, 407.
- [12] D. Bossemeyer, R. A. Engh, V. Kinzel, H. Ponstingl, R. Huber, *EMBO J.* **1993**, *12*, 849.
- [13] C. Kim, N. H. Xuong, S. S. Taylor, *Science* **2005**, *307*, 690.
- [14] R. A. Engh, A. Girod, V. Kinzel, R. Huber, D. Bossemeyer, *J. Biol. Chem.* **1996**, *271*, 26157.
- [15] R. E. London, M. E. Perlman, D. G. Davis, *J. Magn. Reson.* **1992**, *97*, 79.
- [16] B. A. Borgias, M. Gochin, D. J. Kerwood, T. L. James, *Prog. Nucl. Magn. Reson. Spectrosc.* **1990**, *22*, 83.
- [17] J. Zheng, C. B. Post, *J. Magn. Reson. Ser. B* **1993**, *101*, 262.
- [18] T. M. G. Koning, R. Boelens, R. Kaptein, *J. Magn. Reson.* **1990**, *90*, 111.
- [19] N. R. Nirmala, G. M. Lippens, K. Hallenga, *J. Magn. Reson.* **1992**, *100*, 25.
- [20] D. M. Schneider, M. J. Dellwo, A. J. Wand, *Biochemistry* **1992**, *31*, 3645.
- [21] A. L. Lee, P. F. Flynn, A. J. Wand, *J. Am. Chem. Soc.* **1999**, *121*, 2891.
- [22] G. M. Clore, A. M. Gronenborn, *J. Magn. Reson.* **1983**, *53*, 423.
- [23] F. Ni, H. A. Scheraga, *Acc. Chem. Res.* **1994**, *27*, 257.
- [24] T. Langer, M. Vogtherr, B. Elshorst, M. Betz, U. Schieborr, K. Saxena, H. Schwalbe, *ChemBioChem* **2004**, *5*, 1508.
- [25] M. Kroemer, M. K. Dreyer, K. U. Wendt, *Acta Crystallogr. D* **2004**, *60*, 1679.
- [26] W. Kabsch, *J. Appl. Crystallogr.* **1993**, *26*, 795.
- [27] T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr. Sect. A* **1991**, *47* (Part 2), 110.
- [28] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126.
- [29] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53*, 240.
- [30] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65.