

One Crystal, Two Temperatures: Cryoooling Penalties Alter Ligand Binding to Transient Protein Sites

Marcus Fischer,^[b] Brian K. Shoichet,^[b] and James S. Fraser^{*[a]}

This work is dedicated to Professor Thomas Alber.

Interrogating fragment libraries by X-ray crystallography is a powerful strategy for discovering allosteric ligands for protein targets. Cryoooling of crystals should theoretically increase the fraction of occupied binding sites and decrease radiation damage. However, it might also perturb protein conformations that can be accessed at room temperature. Using data from crystals measured consecutively at room temperature and at cryogenic temperature, we found that transient binding sites could be abolished at the cryogenic temperatures employed by standard approaches. Changing the temperature at which the crystallographic data was collected could provide a deliberate perturbation to the equilibrium of protein conformations and help to visualize hidden sites with great potential to allosterically modulate protein function.

Fragment-based ligand discovery (FBLD) uses small-molecule fragments (< 250 Da) to increase the probability of finding weak hits.^[1] Additionally, these fragments have less molecular complexity and can therefore sample chemical space more efficiently than the larger molecules found in conventional high-throughput screening libraries. Successful FBLD campaigns have rapidly improved the affinity of these hits^[2] and advanced several lead molecules to modulate biology or disease.^[3] Fragment-screening methods, including surface plasmon resonance and NMR spectroscopy, can measure binding affinities and provide initial structure–activity relationships (SARs). X-ray crystallography provides key insights for FBLD by identifying the binding sites of hits and structurally guiding medicinal chemistry efforts to optimize the fit (and ultimately the affinity) between the ligand and the binding site.^[4]

Even weak affinity hits can be identified by most FBLD X-ray crystallography experiments, because highly concentrated solutions of ligands are soaked into the protein crystal. The crystal is then cryoooled to protect against radiation damage^[5] and to halt any destructive effects of the solvent or ligands on the

crystal lattice.^[6] Cryoooling also offers tremendous benefits for the transportation of crystals and automated crystal screening by using robotics. An additional, but rarely considered, potential advantage of cryoooling is that thermodynamics favor ligand binding at lower temperatures. For example, given a driving force of 2.16 kcal mol⁻¹ for the standard Gibbs free energy of ligand binding (corresponding to a 26 mM K_d at room temperature) and a concentration of 33 mM for the ligand soaked into the crystal, we would expect the percentage of ligand-bound receptors to be 56% at 293 K (room temperature), 88% at 200 K (the glass-transition temperature), and 99.9% at 100 K (cryogenic temperature; see the Supporting Information). Therefore, at high, but non-saturating, ligand concentrations, greater ligand occupancy at cryoooled temperatures would increase the observable electron density for weak ligands at receptor binding sites.

In addition to active-site ligands, X-ray crystallographic fragment screens often identify hits to secondary binding sites that are distant from the active site and require protein conformational flexibility to become accessible to ligands. Initial hits can be subsequently refined to affect allosteric inhibition or activation, as has been demonstrated against targets such as HIV reverse transcriptase^[7] and 3-phosphoinositide-dependent kinase 1.^[8] These “cryptic” binding sites can be invisible to experimental techniques such as crystallography^[9] because the energy gap between the pocket-forming (high-energy) state and the pocket-occluding (ground) state is too large. Therefore, identifying partially occupied ligands that show only weak electron density is especially important when searching for allosteric modulators, because the electron density will reflect the equilibrium between the pocket-occluding and ligand-bound states. As for active sites of enzymes, cryoooling should increase the fraction of ligands bound at the cryptic sites, which creates a potential advantage for identifying weak cryptic-site binders that can be developed to affect allosteric responses.

However, recent studies suggest that collection of X-ray diffraction data at cryogenic temperatures could mask alternate conformational states that are accessible to the protein at room temperature.^[10] By altering the equilibrium of protein conformations, cryoooling may therefore stabilize the pocket-occluding states of cryptic binding sites and oppose the predicted enhanced ligand occupancy at lower temperatures. We have recently found that considering such high-energy/low-occupancy states, which are present only at room temperature, can be essential for discovering new ligands using flexible receptor docking.^[11] This approach identified ligands that stabi-

[a] Prof. Dr. J. S. Fraser
Department of Bioengineering and Therapeutic Sciences
University of California San Francisco
600 16th St., Genentech Hall, S472E
Box 2240, San Francisco, CA 94158 (USA)
E-mail: james.fraser@ucsf.edu

[b] Dr. M. Fischer, Prof. Dr. B. K. Shoichet
Department of Pharmaceutical Chemistry
University of California San Francisco
1700 4th St., Byers Hall, BH-501, Box 2550, San Francisco, CA 94158 (USA)

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201500196>.

lize specific alternative loop conformations of the cavity site of cytochrome *c* peroxidase (CcP-ga).^[11] Interestingly, in the apo structure (determined at cryogenic temperatures), the conformation that is preferred by the most potent compounds is not significantly populated. These results may also reflect non-equilibrium kinetic considerations of cryocooling,^[12] because cryocooling may occur faster than some protein conformational changes, ligand binding/dissociation events (k_{on}/k_{off}), and ligand diffusion through vitrifying solvent channels. Herein, we have investigated the effect of cryocooling on fragment ligand-binding sites of CcP, one of which is a cryptic site that is only observed upon ligand binding and is not visible in the apo structures. Our studies demonstrate how distinct binding sites can be differentially affected by the tradeoffs between enhancing ligand occupancy upon cryocooling and altering the population of higher energy conformational states that can present new ligand-binding sites.

To systematically probe the impact of temperature on fragment binding, we obtained crystallographic data at both cryogenic temperature and room temperature (RT) from single crystals for ligand-free (apo) CcP-ga and five different ligand-bound CcP-ga complexes (Supporting Information). To minimize the difference between the datasets we 1) used 2-methyl-2,4-pentanediol (MPD) as a precipitant and cryoprotectant to allow for the re-collection of data on the same crystal at cryogenic temperature, 2) collected data on the same crystal volume, 3) matched the crystal size to the synchrotron beam size. This approach minimizes differences in the chemical composition of the crystal, local differences in crystal quality, and differences in the distribution of soaked compounds, respectively. For each crystal, we first measured the diffraction pattern at RT, and then, after flash-cooling the crystal in liquid nitrogen, we re-measured the diffraction pattern of the same crystal at cryogenic temperatures. Changes in the resolution and mosaicity were minimal and typical for well-cooled crystals (Supporting Information, Table S3). Using the same method, we also managed to collect eight complete RT datasets on the same crystal volume, three of which were before the data quality decreased to below 2.1 Å (see also the Supporting discussion and Supporting Information, Figure S8). To compare the structural impact of cryocooling and ligand binding, we calculated the distance of each protein residue from the protein center-of-mass as a function of temperature and as a function of ligand state (bound or apo). The distribution of differences between these distances (e.g., CcP-ligand1 at RT—CcP-ligand1 at cryogenic temperature) provides an estimate of the anisotropy of the structural perturbation of the protein.^[13] Although conformational changes are required to accommodate different ligands,^[11] we found that the protein structure is more perturbed by temperature than by ligand binding (Figure 1A; Figure S1 and Table S8).

For example, the peaks of the distributions comparing datasets at two different temperatures (RT versus cryogenic) with the same ligand state are offset from zero, which shows that there is thermal contraction of the protein. Also, the distributions are broad, which indicates that there is heterogeneity in the structures (Figure 1A). Comparing the apo to the ligand-

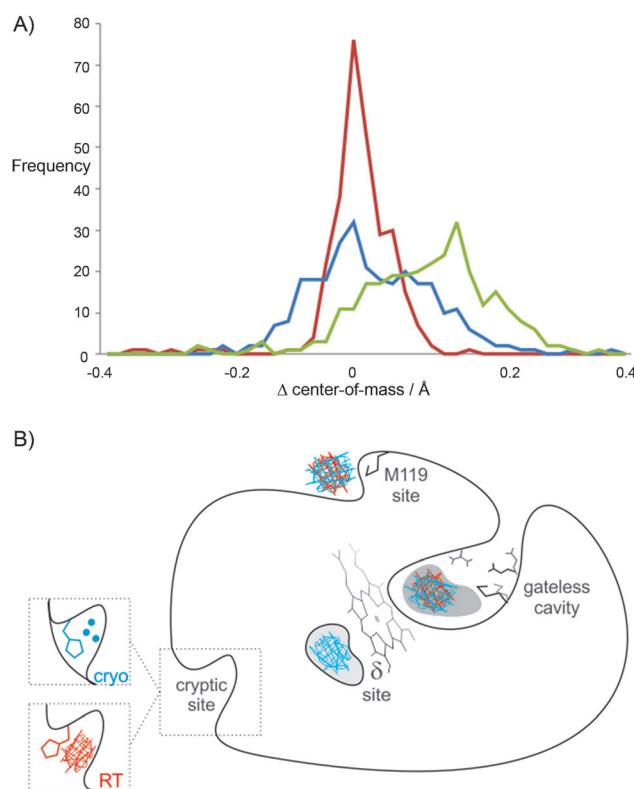


Figure 1. Protein structure is more perturbed by temperature than by ligand binding. A) The distance of each protein residue from the protein center-of-mass is compared between two structures either at different temperatures (green line) or in different ligand states (apo versus with benzimidazole) at the same temperature (red = RT and blue = cryogenic temperature). All temperature pairs were collected consecutively on the same crystal. The amount of offset from zero reflects the expected thermal contraction of the protein upon freezing (green line). The broader distributions indicate structural heterogeneity upon ligand binding at cryogenic temperature (blue line) and structural heterogeneity of the same structure collected at different temperatures (green line). The narrow distribution of the different ligand states at RT (red line) suggests that, at cryogenic temperature, the protein structure is non-specifically perturbed by cryocooling rather than showing a response unique to ligand binding. B) Multiple protein sites display ligand electron density at different temperatures. Electron density was observed at both RT (red mesh) and cryogenic temperature (blue mesh) for the primary cavity site and the M119 surface site, whereas the heme proximal δ-site and the H96 cryptic binding site are temperature sensitive.

bound protein structures at the same temperature reveals that peaks are centered at zero. However, the distributions are much narrower at RT than at cryogenic temperatures, which indicates a decrease in precision at cryogenic temperature, commonly defined as an increase in random errors. Note that the observed decrease in the atomic displacement distributions^[12] (B factors) observed at cryogenic temperatures would normally be interpreted as evidence for increased precision (Tables S1 and S2).

However, the broad distributions of the cryocooled structures suggest much larger structural differences, which are likely due to nonspecific perturbations caused by the cryocooling process and not due to ligand binding. In contrast, the comparisons between the RT structures isolate those structural responses that are unique to ligand binding. This discrepancy between the RT and the cryogenic data is counterintuitive, be-

cause we would expect dissimilarities to be amplified between the ligand-stabilized and the ligand-free structures upon increasing the thermal motion at RT. These results indicate that cryo cooling can have a large and inconsistent impact on the conformations of residues throughout the protein (Figure S2), which may misinform structure-based drug and probe design.

Next, we examined the binding site and other protein regions for large temperature-dependent changes in electron-density distributions. For the ligand benzimidazole, as expected, we observed consistent electron density for the ligand at both temperatures for the primary binding site.^[11] We were surprised to observe ligand density at three additional distal sites (Figure 1B). As with the primary binding site, ligand electron density appears at both temperatures in a second, surface-exposed site near Met119 (Figures 1B; Figure S3B). In a third site, near the δ -heme edge, a known CcP substrate site,^[14] we observed ligand density only at cryogenic temperatures and not at RT (Figures 1B; Figure S3A). Benzimidazole also occupied a fourth site, near His96, but only at RT (Figures 1B and 2). This fourth site can be classified as a cryptic site because the binding pocket is not apparent in the apo structure at either room- or cryogenic temperature. Access of the ligand to this cryptic site is controlled by an alternative conformation of His96, which is correlated with the presence of the benzimidazole (Figure 2) and can be identified by a secondary electron-density peak using Ringer^[15] (Figure 2A). Although the cryogenic data were collected on the same crystal volume as the RT data, the cryogenic temperature electron-density maps are consistent only with the “closed” His96 rotamer and three water molecules occluding the cryptic site (Figure 2B). This result illustrates how the population of alternative conformations can be perturbed by temperature, thus altering the potential for observing a ligand in a binding pocket (Figure 1B).

To probe the biological relevance of the cryptic binding site, we moved from the CcP-ga model system, which has a cavity engineered to bind small molecules,^[11,16] to CcP-wt, which contains the radical-forming Trp191 residue that is involved in long-range electron transfer at the active site.^[17] We first confirmed that benzimidazole occupied the cryptic binding site in CcP-wt. Co-crystals with benzimidazole diffracted to a resolution of 2.6 Å and showed electron density for the ligand in the cryptic binding site for three of the four protein copies within the crystallographic asymmetric unit.

Refinement of the data unambiguously revealed His96 to be in an open state (Figure S4), but no ligand density at the Met119 residue or δ -site could be detected at this resolution. To determine the binding affinity of benzimidazole to the cryptic binding site, we monitored ligand binding through the saturable perturbation of the CcP heme Soret band.^[16b] In the CcP-ga cavity site, benzimidazole can occupy multiple sites and has low micromolar affinity for the primary binding site (Figure S5). However, in the wild-type protein, Trp191 blocks this high-affinity binding site, which allowed us to isolate the binding affinity of benzimidazole to the lower affinity site. We determined an affinity of benzimidazole for CcP-wt of 26 mM (Figure S6), which corresponds to a ligand efficiency (LE)^[18] of

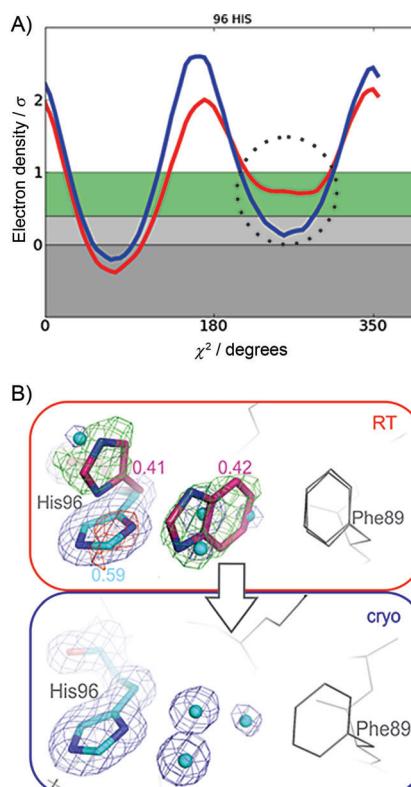


Figure 2. Benzimidazole binding to the transient cryptic site is only observed at RT. **A)** Electron density sampling around His96 Chi-2 using Ringer reveals an electron density peak (dotted circle) for the alternative conformation at room temperature (red line), but not cryogenic temperature (blue line). This minor “open” state would have been missed if the conventional 1σ cutoff to distinguish electron density signal from noise was used or only cryogenic data were available. **b)** Difference electron density contoured at 3σ (green and red mesh) confirms the presence of the alternative His rotamer and ligand presence at RT (red box). Both the ligand and the alternative His96 conformer were excluded from the crystallographic refinement but are shown in magenta as refined automatically to their final occupancies. At cryogenic temperature (blue box) we observed no ligand electron density but instead only the pocket-filling His96 rotamer and water molecules (blue spheres) that occluded the cryptic site. Note that observing ligand binding to the cryptic site only at RT is unexpected, because thermodynamics should favor binding at cryogenic temperatures.

0.24 kcal mol⁻¹ per ligand heavy atom (HA; Figure 2A). Given that the cryptic site is too far away (25 Å) from the heme to elicit a direct Soret shift, we were initially surprised to observe an allosteric Soret band shift for benzimidazole. Although we detected no steric coupling networks connecting the cryptic binding site to the heme by using CONTACT^[19] in our CcP-ga crystal structures at RT (Figure S7), a recent EPR study proposes the cryptic binding site to be a biologically relevant site for substrate oxidation that is remote from the heme.^[20] This alternative electron transfer pathway includes the nearby Tyr71 residue as a reactive intermediate^[20] and contrasts with a model where small-molecule binding to sites other than the δ -heme edge was suggested to be nonspecific.^[21]

The discrepancy between a binding site being occupied by benzimidazole at RT and unoccupied in the same crystal at cryogenic temperature is counterintuitive and suggests that cryo cooling can “overwhelm” the driving forces of ligand bind-

ing. To explain the reduction of benzimidazole occupancy from approximately 50% at room temperature to below the detection limit of about 5% at cryogenic temperature, cryo cooling must counteract the temperature-dependent effects on the ligand–protein equilibrium. To estimate the penalty on occupancy from cryo cooling, we assume both cryo cooling and ligand soaking to be at equilibrium, although the heterogeneous structural perturbations we observed in response to cryo cooling could suggest that the system has not equilibrated at the cryo cooled temperature. However, on the typical timescale of freezing (100 ms),^[12] we are neither in the fast-cooling regime where RT states are frozen-in nor in the slow-cooling regime where the crystal has had time to equilibrate. Therefore, intermediate states of slow and fast equilibrating domains are likely differentially trapped. We related the energy penalty (ΔG) as a function of ligand-soaking concentration [L], standard affinity at room temperature ($K_{d\text{site}}$), minimum detectable occupancy (occ_{\min}), temperature T , and the gas constant R :

$$\Delta G_{\text{penalty}} > -RT \ln \frac{[(1-\text{occ}_{\min})/\text{occ}_{\min}][L]}{K_{d\text{site}}} \quad (1)$$

See Supporting Information for the derivation of this equation.

Given our experimental conditions of soaking a ligand with a K_d of 26 mM at a concentration of 33 mM and crossing the glass-transition temperature of 200 K upon cooling, the allosteric cryo cooling penalty must be at least 1.3 kcal mol⁻¹ to render the ligand invisible below 5% occupancy. The cryo cooling penalty includes the solvent glass transition, unusual temperature dependencies of other enthalpic or entropic terms, and allosteric lattice changes to the protein ensemble. All of these mechanisms are likely to be important for increasing the magnitude of the cryo cooling penalty of CcP residue His96.^[12]

In contrast to the behavior of benzimidazole at the cryptic binding site, we identified two examples where the cryo cooling penalty does not dominate over other contributions to binding. First, 2-amino-5-methylthiazole binds to the cryptic binding site at both temperatures (Figure 3 A). When the soaking concentration is 100 mM, 2-amino-5-methylthiazole ($k_d = 68$ mM; LE = 0.22) shows higher crystallographic occupancy at cryogenic temperatures than at RT (75% versus 54%, respectively; Figure 3 A); this mirrors our expectation of 99.6% occupancy at 100 K and 59% occupancy at 298 K. Second, cryo cooling can promote compounds to bind to other sites that do not require a conformational change to form a pocket. For the δ -site, for example, we only observed benzimidazole density at cryogenic temperatures (Figures 1B and S3). Based on these results, which suggest that the cryo cooling penalty would disfavor binding at the cryptic site, and the fact that the Protein Data Bank is dominated by data collected at cryogenic temperatures,^[5] we expected to find more δ -site binders than cryptic site binders in the previously determined structures. We inspected 136 electron-density maps of previously determined structures of three CcP variants (CcP-wt, CcP-W191G, CcP-ga) from the PDB and struggled to find significant electron density in the δ -site, whereas we found several examples of ligand densities in the cryptic binding site (Figure 3 B). Counterintui-

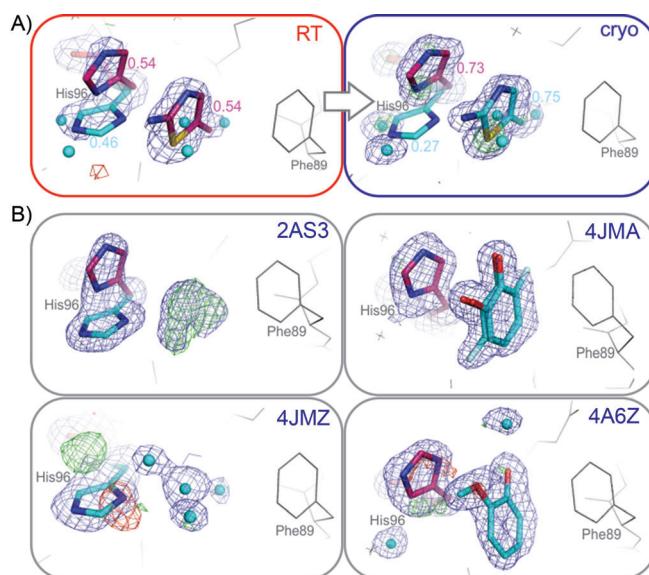


Figure 3. Cryptic site binders are prevalent in the PDB. A) 2-amino-5-methylthiazole binds to the cryptic site at both temperatures, RT (red box) and cryogenic (blue box). Both the ligand and the alternative His96 conformation (included in refinement) refined to higher ligand occupancy at cryogenic temperatures (0.75 versus 0.54 at RT). B) Electron density maps in other CCP structures, as deposited in the PDB, show evidence of ligand binding with an “open” His96 conformer, either unmodeled (phenol in PDB structure 2AS3, and N-methyl-1H-benzimidazol-2-amine in 4JMZ) or modeled (3-fluorocatechol in 4JMA, and guaiacol in 4A6Z). 2mFo-DFc maps shown as blue mesh (rendered at 1 σ), mFo-DFc maps in green and red (3 σ).

tively, this suggests a lower driving force for ligands to bind to the δ -site. Because the δ -site is accessible without local side-chain conformational changes, the lower observed frequency might also hint at allosteric lattice effects that reduce binding at this site. Collectively, these results illustrate how cryo cooling can have counteracting effects on ligand occupancy at fragment-binding sites.

Heterogeneous and non-equilibrium contributions to protein–ligand interactions upon cryo cooling make it difficult to assign an exact value for this penalty beyond our previously mentioned estimate of 1–2 kcal mol⁻¹. However, our observations illustrate that the cryo cooling penalty plays a dominant role in determining the net binding of a ligand in a cryogenically frozen protein.

Our observation that the cryptic binding site was occupied at RT but not at cryogenic temperature contradicts the thermodynamic expectation that a higher fraction of sites should be bound at cryogenic temperatures. This suggests shifting temperatures as a general strategy to modulate the energy landscape of protein–ligand binding and overcome cryo cooling penalties in favor of populating and revealing transient sites. We note that these cryo cooling effects can be especially problematic at ligand concentrations around or below the K_d , which equals the ligand concentration at which half of the protein molecules are bound. Although the observation of differential binding depends on a lucky choice of concentration and compound soaking time, as for benzimidazole in CcP-ga, the chance of observing a secondary ligand-binding site increases with concentration—at both temperatures. Fragments will be

particularly affected, because they intrinsically achieve low binding affinities even at high ligand efficiencies, and the penalties may therefore overwhelm binding upon cryo cooling. To counteract this effect, very high soaking concentrations would have to be used to achieve a sufficient fraction of receptors bound to a ligand. However, preparing such high concentration stock solutions is often impractical, because it is limited by the solubility of compound and the sensitivity of the protein to organic solvents (like DMSO) or the compound itself.

Allosteric ligand-binding sites offer great potential for modulating protein function, but are often difficult to visualize. However, cryptic binding sites, with the potential to allosterically modulate protein function, can be discovered serendipitously, even for well-studied proteins, by using a fragment-based approach. Herein, we demonstrate that shifting the temperature at which the crystallographic data are collected can deliberately perturb the protein to help visualize such cryptic binding sites. To shift the population of conformational states towards the energetically less-accessible states and detect new binding sites for low-affinity, less-soluble fragments, we suggest a dual strategy: collect both datasets, if possible. This approach will complement mutagenesis efforts designed to stabilize specific protein conformations and may help identify cryptic binding sites, which could substantially extend the targets that can be probed to dissect biological mechanisms or enable therapeutic intervention.^[9,22] For fragments, rather than invalidating cryogenic data, RT data collection has potential as an orthogonal method that could unleash some of the unused potential within FBLD. However, even without RT data, we suspect that many existing electron-density maps may contain evidence of unmodeled ligands partially occupying such sites; especially when high concentrations of ligand were used to soak the crystals, as is typical of fragment-based methods. While we cannot link remote binding sites to function from structure alone, FBLD-investigators specifically looking to find or optimize allosteric binding ligands may find promise in exploring the full landscape of ligand binding. In those cases, our strategy of shifting conformational equilibria by shifting temperature may become illuminating, even crucial.

Experimental Section

Experimental details are given in the Supporting Information.

Acknowledgements

We thank Allison Doak for help with the protein preparation and Dr. Anja Fischer, Dr. Christoph Rademacher, and Chelsea Bidlow for critical reading of this manuscript. M.F. and B.K.S. are supported by GM59957. J.S.F. is a Searle Scholar, Pew Scholar, and Pack-

ard Fellow, and is supported by grants: NIH OD009180, NIH GM110580, and NSF STC-1231306. Data collection at BL831 at the Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and the Program Breakthrough Biomedical Research, which is partially funded by the Sandler Foundation.

Keywords: allosterism • biophysics • ligand discovery • structural biology • thermodynamics • X-ray diffraction

- [1] M. Fischer, R. E. Hubbard, *Mol. Interventions* **2009**, *9*, 22–30.
- [2] P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **1997**, *278*, 497–499.
- [3] a) C. W. Murray, M. L. Verdonk, D. C. Rees, *Trends Pharmacol. Sci.* **2012**, *33*, 224–232; b) C. Lipinski, A. Hopkins, *Nature* **2004**, *432*, 855–861.
- [4] D. Patel, J. D. Bauman, E. Arnold, *Prog. Biophys. Mol. Biol.* **2014**, *116*, 92–100.
- [5] E. Garman, *Curr. Opin. Struct. Biol.* **2003**, *13*, 545–551.
- [6] T. G. Davies, I. J. Tickle, *Top. Curr. Chem.* **2012**, *317*, 33–59.
- [7] J. D. Bauman, D. Patel, C. Dharia, M. W. Fromer, S. Ahmed, Y. Frenkel, R. S. Vijayan, J. T. Eck, W. C. Ho, K. Das, A. J. Shatkin, E. Arnold, *J. Med. Chem.* **2013**, *56*, 2738–2746.
- [8] J. D. Sadowsky, M. A. Burlingame, D. W. Wolan, C. L. McClendon, M. P. Jacobson, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6056–6061.
- [9] G. R. Bowman, P. L. Geissler, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11681–11686.
- [10] a) J. S. Fraser, M. W. Clarkson, S. C. Degnan, R. Erion, D. Kern, T. Alber, *Nature* **2009**, *462*, 669–673; b) J. S. Fraser, H. van den Bedem, A. J. Samelson, P. T. Lang, J. M. Holton, N. Echols, T. Alber, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 16247–16252.
- [11] M. Fischer, R. G. Coleman, J. S. Fraser, B. K. Shoichet, *Nat. Chem.* **2014**, *6*, 575–583.
- [12] B. Halle, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4793–4798.
- [13] a) S. D. Rader, D. A. Agard, *Protein Sci.* **1997**, *6*, 1375–1386; b) R. F. Tilton, Jr., J. C. Dewan, G. A. Petsko, *Biochemistry* **1992**, *31*, 2469–2481.
- [14] A. Gumiero, E. J. Murphy, C. L. Metcalfe, P. C. Moody, E. L. Raven, *Arch. Biochem. Biophys.* **2010**, *500*, 13–20.
- [15] P. T. Lang, H.-L. Ng, J. S. Fraser, J. E. Corn, N. Echols, M. Sales, J. M. Holton, T. Alber, *Protein Sci.* **2010**, *19*, 1420–1431.
- [16] a) S. Barelier, S. E. Boyce, I. Fish, M. Fischer, D. B. Goodin, B. K. Shoichet, *PLoS One* **2013**, *8*, e69153; b) G. J. Rocklin, S. E. Boyce, M. Fischer, I. Fish, D. L. Mobley, B. K. Shoichet, K. A. Dill, *J. Mol. Biol.* **2013**, *425*, 4569–4583.
- [17] A. N. Volkov, P. Nicholls, J. A. Worrall, *Biochim. Biophys. Acta Bioenerg.* **2011**, *1807*, 1482–1503.
- [18] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discovery Today* **2004**, *9*, 430–431.
- [19] H. van den Bedem, G. Bhabha, K. Yang, P. E. Wright, J. S. Fraser, *Nat. Methods* **2013**, *10*, 896–902.
- [20] K. D. Miner, T. D. Pfister, P. Hosseinzadeh, N. Karaduman, L. J. Donald, P. C. Loewen, Y. Lu, A. Ivancich, *Biochemistry* **2014**, *53*, 3781–3789.
- [21] E. J. Murphy, C. L. Metcalfe, C. Nnamchi, P. C. Moody, E. L. Raven, *FEBS J.* **2012**, *279*, 1632–1639.
- [22] A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug. Discov.* **2002**, *1*, 727–730.

Manuscript received: April 16, 2015

Accepted article published: May 28, 2015

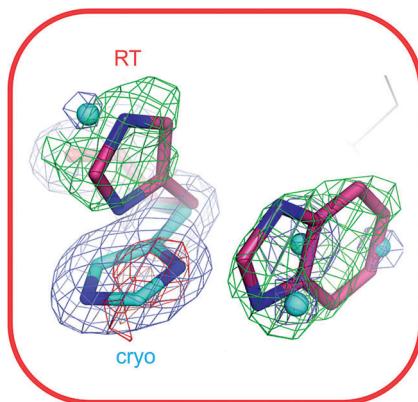
Final article published: ■■■, 0000

COMMUNICATIONS

M. Fischer, B. K. Shoichet, J. S. Fraser*

■ ■ - ■ ■

 One Crystal, Two Temperatures:
Cryocooling Penalties Alter Ligand
Binding to Transient Protein Sites



Too cool to show up: The common practice of cryocooling crystals was found to mask transient protein conformations. Data collected on a single crystal revealed ligand binding to a cryptic site at room temperature that was hidden at cryogenic temperature. These results contradict thermodynamic expectations and provide a method for discovering allosteric sites with great potential for modulating protein (mal)-function.

CHEMBIO**CHEM**

Supporting Information

One Crystal, Two Temperatures: Cryoooling Penalties Alter Ligand Binding to Transient Protein Sites

Marcus Fischer,^[b] Brian K. Shoichet,^[b] and James S. Fraser^{*[a]}

cbic_201500196_sm_misellaneous_information.pdf

Supplemental Methods

Derivation of equation 1:

$$(1) \Delta G_{app} = \Delta G_{site} + \Delta G_{penalty}$$

$$(2) fraction_{bound} = \frac{[L]/Kd_{app}}{1+[L]/Kd_{app}}$$

$$-RT \ln Kd_{app} = -RT \ln Kd_{site} + \Delta G_{penalty}$$

$$Kd_{app} = e^{\ln Kd_{site} + (\Delta G_{penalty}/-RT)} = e^x$$

Assuming a minimum occupancy of 5% to observe binding

$$occ_{min} > \frac{\frac{[L]}{e^x}}{1+\frac{[L]}{e^x}} \quad (\text{equation 2})$$

$$0.05 > \frac{\frac{[L]}{e^x}}{1 + \frac{[L]}{e^x}}$$

$$0.05 + \frac{0.05 [L]}{e^x} > \frac{[L]}{e^x}$$

$$0.05 e^x + 0.05 [L] > [L]$$

$$e^x > 19 [L]$$

$$x > \ln(19 [L])$$

$$\ln Kd_{site} + (\Delta G_{penalty}/-RT) > \ln(19 [L])$$

$$\Delta G_{penalty} > -RT \ln \frac{19 [L]}{Kd_{site}}$$

Generalization: $\Delta G_{penalty} > -RT \ln \frac{\frac{(1-occ_{min})}{occ_{min}} [L]}{Kd_{site}}$ (equation 1)

Calculation of temperature-dependence of occupancies

The standard Gibbs energetic driving force for benzimidazole binding is 2.16 kcal/ mol at room temperature. We first calculate Kd_{app} from $\Delta G = -RT \ln Kd_{app}$ for each temperature and then determine the occupancies via equation 2 as:

	RT	200K	100K
[L] (M)	0.033	0.033	0.033
Kd_{app} (M)	0.026	0.00435	0.0000189
occupancy	0.559	0.884	0.999

Protein purification and crystallography

The protein was purified and crystallized as described¹. Crystals were soaked at RT at concentrations indicated in the pdb file with compounds directly dissolved in the cryoprotectant 25% MPD. To minimize differences between the datasets a) MPD (2-methyl-2,4-pentanediol) was used as a precipitant and cryoprotectant to allow immediate re-collection of data on the same crystal at cryogenic temperature, b) data were collected on the same crystal volume, which matched the beam size. Matching the crystal to the beam size and using the same spot at both temperatures should keep any differences in the local distribution of soaked compounds from the rim to the core of the center of the crystal constant between both datasets. Therefore differences due to this effect can be excluded. Data were collected at the ALS Berkeley beamline 8.3.1 and automatically processed using the xia2 pipeline ² to ensure that unit cell dimensions are

determined individually. Structures were solved by molecular replacement with Phaser³ using the same model and R_{free} for all structures. Alternating cycles of refinement and model building were carried out in Refmac⁴ or PHENIX⁵, and Coot⁶. Alternative conformations and the respective ligand were added at the late stages of refinement. Phenix.refine⁵ was used for occupancy refinement (strategy=individual_adp+occupancies). RT models were collected on the same crystal as previously deposited structures collected at cryo and deposited to the PDB as listed in SI Table 1. The wt structure was solved using the AutoMR and AutoBuild functions in PHENIX without manual intervention to avoid bias. There were two lines of evidence for ligand binding: The His96 appeared in the open conformation and there was density for the ligand. Therefore the ligand, benzimidazole, was included in the final stage of the occupancy refinement using phenix.refine (deposited to the pdb as 4XVA). Prior to deposition all structures were validated using validation tools within Coot, PHENIX, CCP4 and the ADIT deposition server.

We used temperature as a perturbation to modulate the state of the protein – and temperature affects related parameters like mosaicity and resolution. For the temperature pairs, consistent with previous studies⁷⁻⁹, the unit cell size shrinks (by 8%) and mosaicity increases upon cryocooling (SI Table 3). Mosaicity results from differences in unit cell parameters (likely in small blocks within the crystal) and generally increases during cryocooling. Our cryocooled datasets have low mosaicity values that are indeed slightly larger than the RT datasets (SI Table 3). This suggests that our cryocooling procedure has not disrupted the crystal quality. Additionally, the resolution is slightly higher at cryogenic temperatures, likely reflecting the reduced thermal motion. Moreover, neither the slight increases in mosaicity nor gains in resolution should significantly alter the ability resolve the contents of the unit cell.

Despite the best efforts to unify the freezing protocol, these differences are a generally unavoidable phenomenon of cryo crystallography and therefore representative of cryogenic structures in the pdb, which comprise ~95% of the pdb. Experimental differences are rarely documented and hard to account for, even in systematic studies. Differences we considered include: differences in sample and loop size, residual liquid around the crystal, crystal age, temperature and composition, changes in the relative humidity, cooling rate, thickness of cold gas layer above LN₂ dewar, the choice of cooling liquid and the velocity of gas stream. As we ultimately care about finding ligands, awareness of such confounding factors can help to plan experiments and motivate careful analysis, especially if the outcome is (thermodynamically) unexpected and not entirely understood.

When applying the same stringency in the crystallographic refinement, the final model had a significantly lower average R_{free} of 0.14 at RT compared to 0.17 at cryogenic temperatures. The ~3% difference in R_{free} indicates superior model quality at RT although data, on average, were collected to similar resolution and showed no obvious signs of radiation damage, judged by an average R_{merge} of <6% at either temperature (SI Table 1-3).

Electron density sampling to reveal alternative side-chain conformations was performed using Ringer¹⁰. CONTACT¹¹ was used to identify dynamic contact networks between the cryptic site and the heme. Center of mass (C.O.M.) calculations were analogous to Frauenfelder et al.⁸, using <http://bioinformatica.isa.cnr.it/CALCOM/>¹² and Microsoft Excel for plotting. All structural figures were generated in PyMOL.

Binding assays

Affinities were measured by monitoring the shift of the heme Soret band near 410nm as described previously^{1,13}. Ligand binding was measured by endpoint titration in

100 mM citric acid buffer adjusted to pH 4.5 with Bis-Tris propane. Ligand stocks were made up to 1 M in dimethyl sulfoxide.

Supplemental Results and Discussion

Multiple complete, high-resolution datasets can be collected on a single crystal at room temperature.

It is commonly believed that multiple crystals are required to collect one complete dataset at RT. To show that this is not necessarily true we used a single crystal to collect multiple datasets at RT on a third generation synchrotron source. To facilitate consecutive data collection on a single crystal we attenuated the x-ray beam by using an Al foil and shifting the energy to 13000 eV to reduce flux. To avoid evaporation, we covered the crystal with a MiTeGen plastic sleeve containing 10 μ l of 75% reservoir solution and 25% water. We collected eight consecutive datasets on the same crystal volume, with three consecutive datasets collected to a resolution of <2.1 \AA and 7 consecutive datasets collected to a resolution of <3 \AA ; when insisting on >99% completeness, high multiplicity, and $I/\sigma I > 2$ for the highest resolution shell (SI Fig. 8A). This illustrates that multiple datasets can be collected on a single crystal at RT but we note that the number may vary for lower symmetry space groups than P212121. However, CcP contains a radiation-sensitive heme iron, which we expected to impede data quality. We observe similar Fe difference densities throughout collecting 8 redundant datasets and, using Ringer, we found no differences in side-chain conformations of nearby residues His175, that coordinates the iron, and Asp233, which is responsible for binding primarily cationic ligands (SI Fig. 8B) To investigate which residues are most affected in the binding site we use local density correlation (LDC) to compare the electron density distribution at each residue between

two datasets (SI Fig. 8C). Calculating the correlation coefficient between the first and all subsequent datasets show high correlation of local electron densities across the protein. Globally, the correlation coefficients (CC) between datasets decreased continuously from 95.5% to 64.6%, while CC standard deviations increased from 1.3% to 7.2% (SI Fig. 8C). Both trends indicate increasing structural heterogeneity with progressing damage from the first consecutive to the final dataset. Locally, we further looked at residues with proximity to the heme iron, functional relevance, higher propensity to suffer from radiation damage, or localization on the protein. None of these residues showed a noteworthy deviation from the average LDC. In the final dataset the three most affected residues (with the lowest CC) are V45, G41 and V169, and the least affected (with the highest CC) are S225, F274 and E35; all three are on the exterior of the protein (SI Fig. 8C).

Supplemental Figures and Tables

SI Table 1 Overview of PDB structure pairs collected on the same crystals at different temperatures.

ligand ZINC-ID	cryo	RT
(apo)	4OQ7	4NVA
8652421	4NVB	4XV4
36634	4NVC	4XV8
331945	4NVD	4XV7
331902	4NVE	4XV5
1583444	4NVF	4XV6

SI Table 2 Data collection and refinement statistics

PDB ID	4XV5	4XV4	4XVB	4XV7
Ligand	Benzimidazole	2-Amino-5-methylthiazole	Benzamidine	4-Amino-quinazoline
ZINC ID	331902	8652421	36634	331945
Data collection				
Space group	P212121	P212121	P212121	P212121
Cell dimensions				
a, b, c (Å)	51.53, 76.83, 107.57	51.53, 76.52, 107.30	51.46, 76.95, 107.71	51.48, 76.76, 107.69
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	39.76-1.65 (1.70- 1.65)	51.53-1.69 (1.73- 1.69)	39.76-1.57 (1.61- 1.57)	39.74-1.62 (1.66- 1.62)
R_{merge}	0.054 (0.692)	0.068 (0.707)	0.053 (0.555)	0.062 (0.695)
$I / \sigma I$	17.9 (2.1)	13.7 (2.1)	14.9 (2.4)	13.3 (2.0)
Completeness (%)	99.7 (99.9)	99.9 (99.9)	99.6 (99.8)	99.9 (100)
Redundancy	4.2 (4.2)	4.1 (4.2)	4.3 (4.2)	4.3 (4.2)
Mosaicity	0.109	0.153	0.171	0.276
Wilson B	19.3	19.7	19.1	19.1
Refinement				
Resolution (Å)	39.76-1.65	51.53-1.69	39.76-1.57	39.74-1.62
No. reflections	51656 (5113)	48203 (4753)	60162 (5926)	54844 (5385)
$R_{\text{work}} / R_{\text{free}}$	0.1125 / 0.1503	0.1099 / 0.1495	0.1082 / 0.1399	0.1150 / 0.1477
No. atoms				
Protein	2558	2549	2553	2500
Ligand/ion	18	14	9	11
Water	253	268	298	262
B -factors				
Protein	22.0	21.1	21.4	22.3
Ligand/ion	29.5	26.2	14.8	17.2
Water	38.8	39.4	41.3	38.4
R.m.s. deviations				
Bond lengths (Å)	0.018	0.018	0.017	0.017
Bond angles (°)	1.49	1.54	1.59	1.46

*Values in parentheses are for highest-resolution shell.

PDB ID	4XV6	4XVA
Ligand	MES	Benzimidazole/ wtCcP complex
ZINC ID	1583444	331902
Data collection		
Space group	P212121	P212121
Cell dimensions <i>a, b, c</i> (Å)	51.48, 76.52, 107.3	84.11, 104.76, 185.25
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	33.42-1.55 (1.59- 1.55)	92.63-2.66 (2.73- 2.66)
R_{merge}	0.04 (0.65)	0.08 (0.70)
$I / \sigma I$	17.0 (2.2)	12.3 (2.1)
Completeness (%)	99.7 (99.7)	99.6 (97.9)
Redundancy	4.0 (4.1)	4.1 (4.0)
Mosaicity	0.237	0.258
Wilson B	19.9	64.5
Refinement		
Resolution (Å)	33.42-1.55	92.63-2.66
No. reflections	62014 (6158)	47484 (4699)
$R_{\text{work}} / R_{\text{free}}$	0.1035 / 0.1355	0.2326 / 0.2764
No. atoms		
Protein	2543	9464
Ligand/ion	12	208
Water	272	284
<i>B</i> -factors		
Protein	24.4	54.4
Ligand/ion	21.8	44.5
Water	43.4	45.0
R.m.s. deviations		
Bond lengths (Å)	0.019	0.016
Bond angles (°)	1.65	1.39

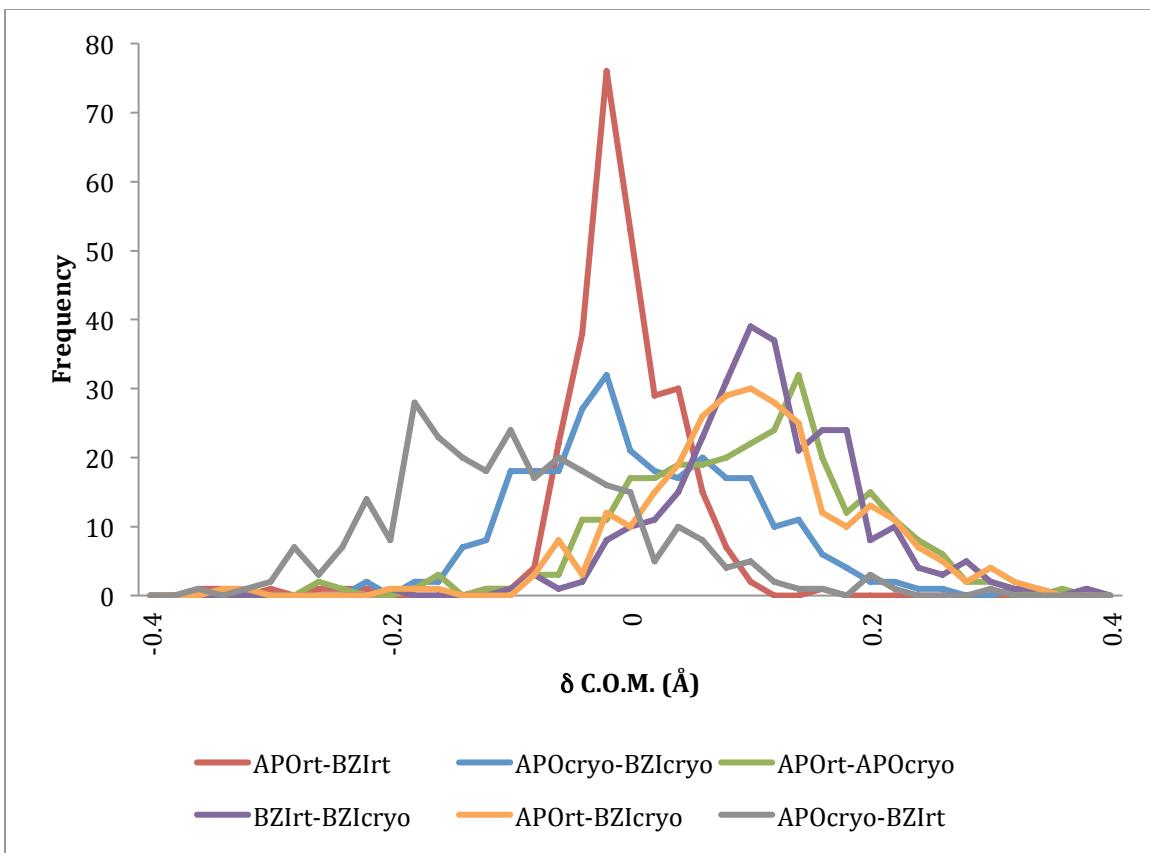
SI Table 3 (a) Values for benzimidazole are comparable at both temperatures. (b) Standard deviations (bold) are larger for cryo data collected on the same crystal; averaged over 6 crystals from SI Table S1.

a)

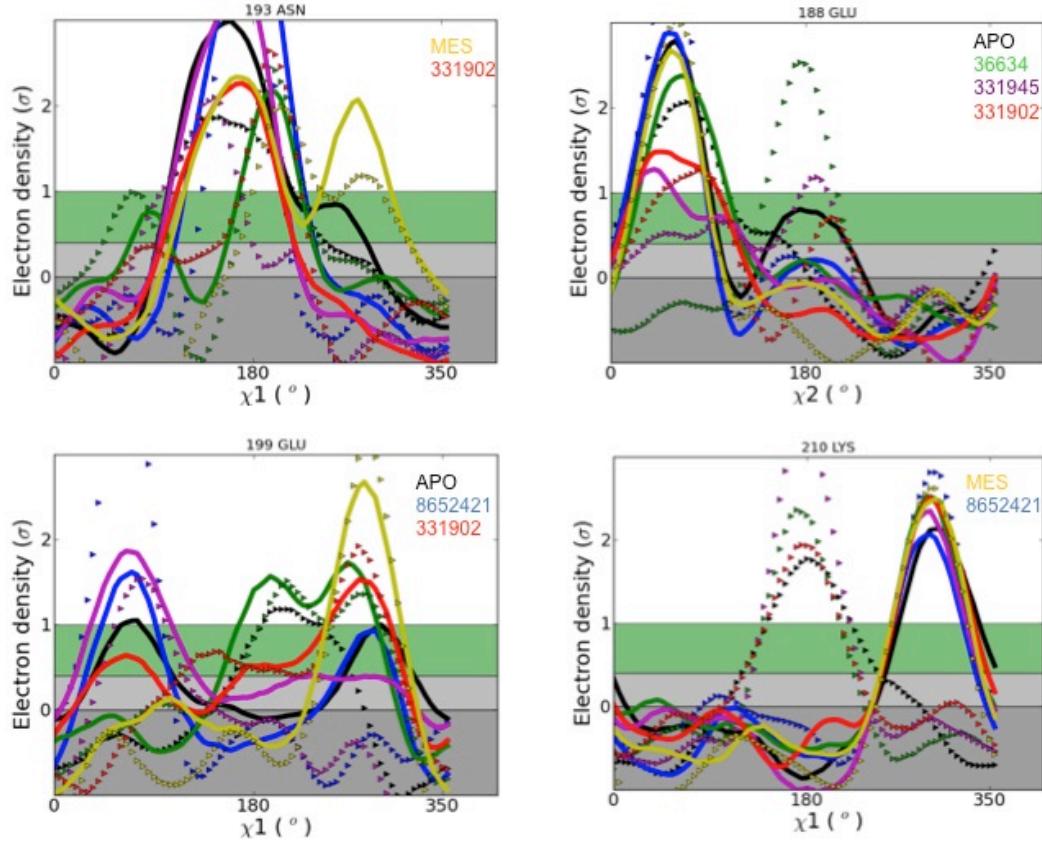
benzimidazole	RT	cryo
resolution	1.65 Å	1.54 Å
R _{merge}	0.054	0.042
mosaicity	0.109	0.175
final R _{free}	0.150	0.180

b)

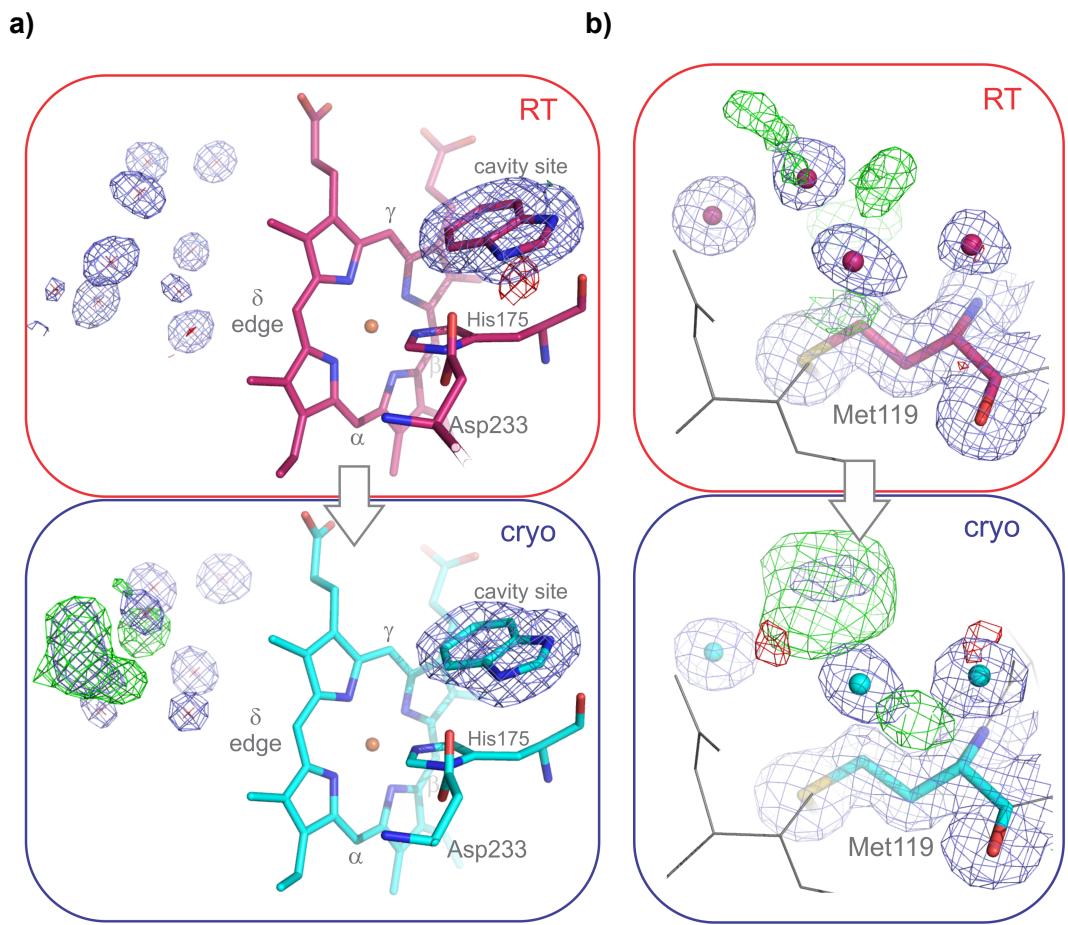
Average	Res	mosaicity	Wilson B	Rfree	UC vol	a	b	c	Protein vol
RT	1.61	0.17	19.4	0.146	424481 ±1678	51.49 ±0.03	76.68 ±0.19	107.5 ±0.18	50078 ±417
cryo	1.42	0.28	11.7	0.169	395543 ±6702	50.83 ±0.11	74.00 ±0.49	105.15 ±1.04	54403 ±3331



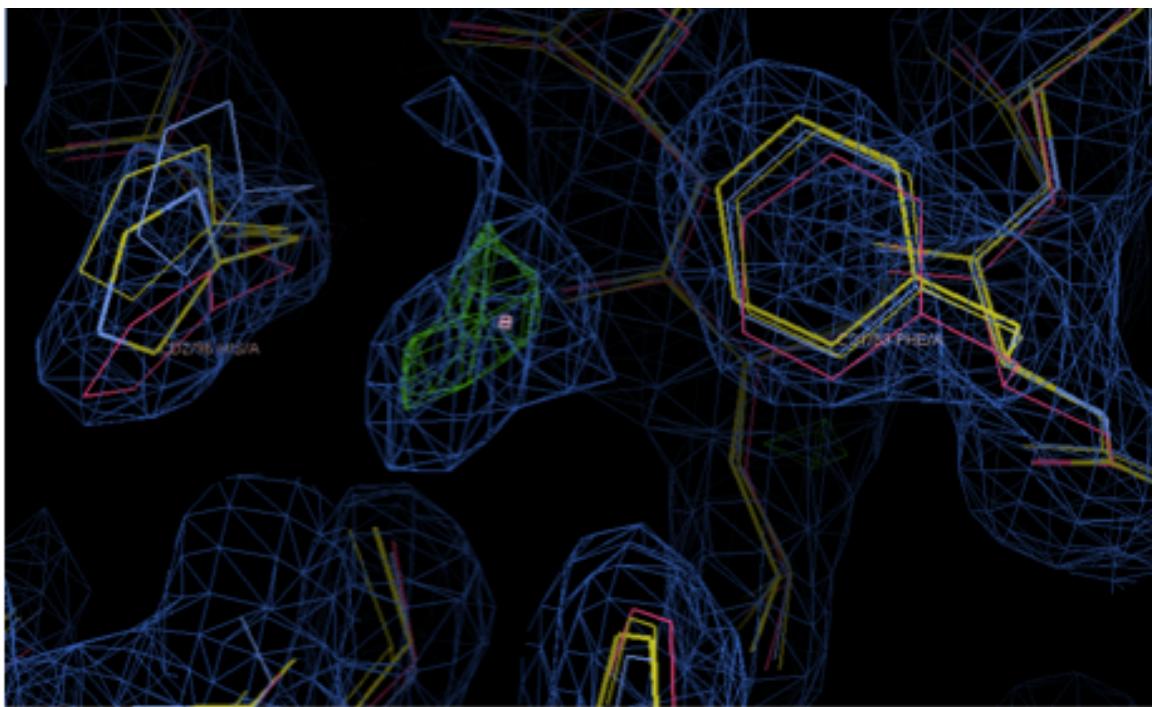
SI Figure 1 As main text Figure 1a with all other combinations.



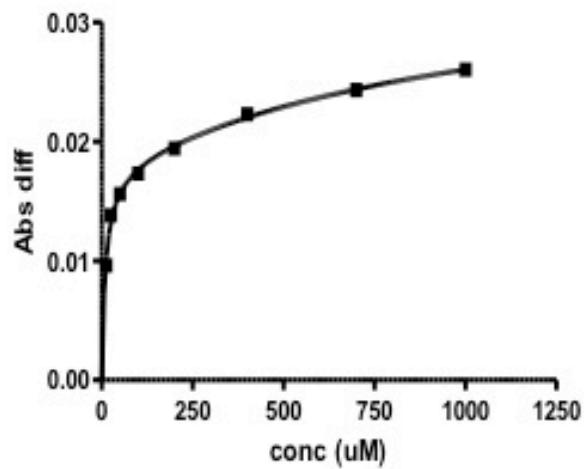
SI Figure 2 Non-systematic changes for several flexible side-chains in response to temperature. Ringer plots show differences in side-chain conformations (chi angles as a function of electron density) when datasets were collected on the same crystal at two temperatures. Temperature pairs of the same complex (or apo structure) are matched in color, with RT data shown as a continuous line and cryogenic data shown as a dashed line. Labels in the top right hand side corner of each plot denote complexes where the corresponding residue conformation responds non-systematically to temperature.



SI Figure 3 Electron density evidence for ligand binding to a) the delta site and the primary cavity site, and b) the M119 site; data of the benzimidazole-CcPga complex were collected on the same crystal at room-temperature (red box) and cryogenic temperature (blue).

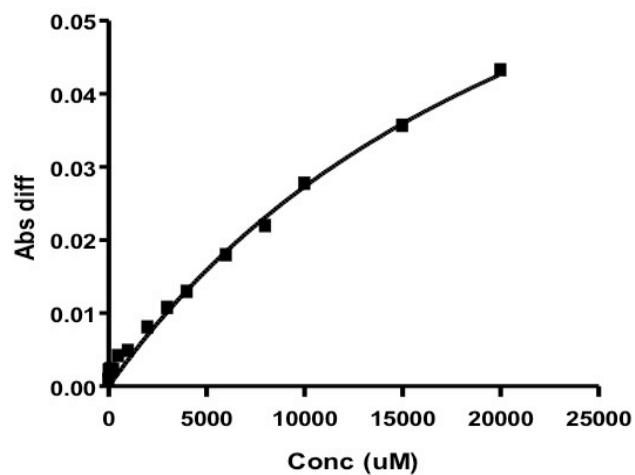


SI Figure 4 Automatic refinement of wildtype CcP data reveals His96 (left) in conformations that open the cryptic site for benzimidazole binding; difference electron density appears with the ligand excluded from refinement.



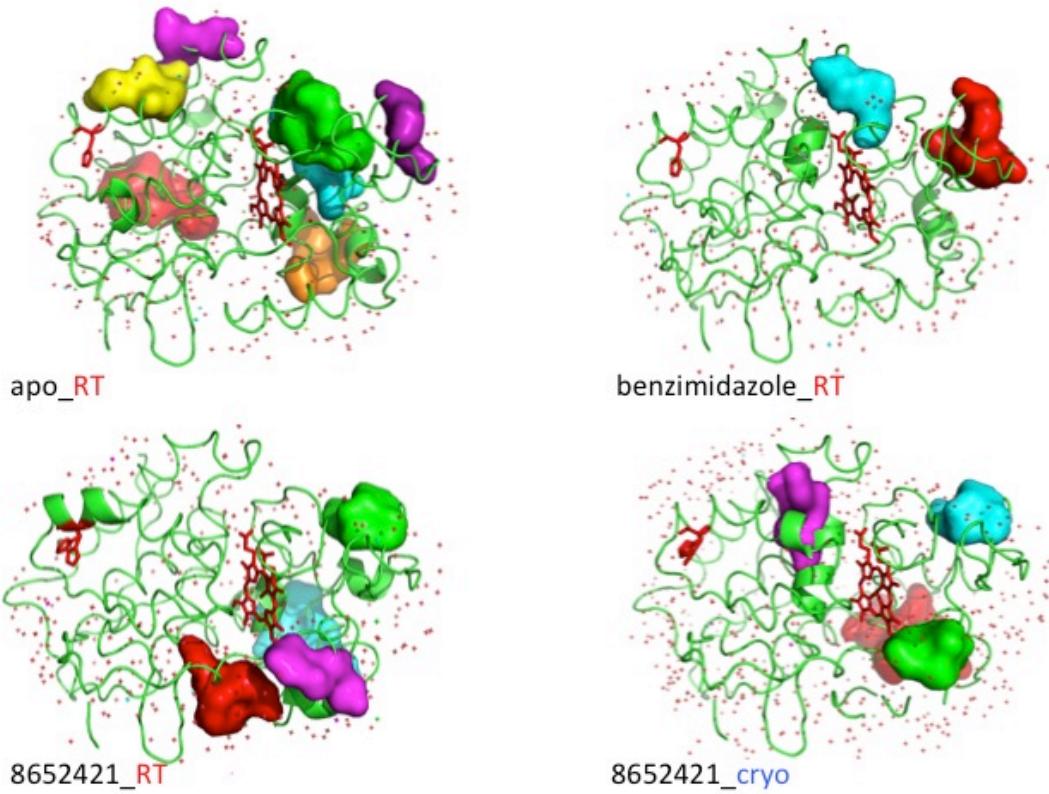
SI Figure 5 Soret band affinity data for CcP cavity mutant and benzimidazole.

A two-site binding model fits a k_{d1} of 10 μM for the cavity site and k_{d2} of >1 mM (1169 μM) for the secondary site. The R^2 for the two-site binding hyperbola is 0.9955 compared to 0.9037 for the single site fit, where only the k_{d1} for the primary site was fitted as 23.5 μM .



SI Figure 6 Soret band affinity data for CcP wildtype and benzimidazole.

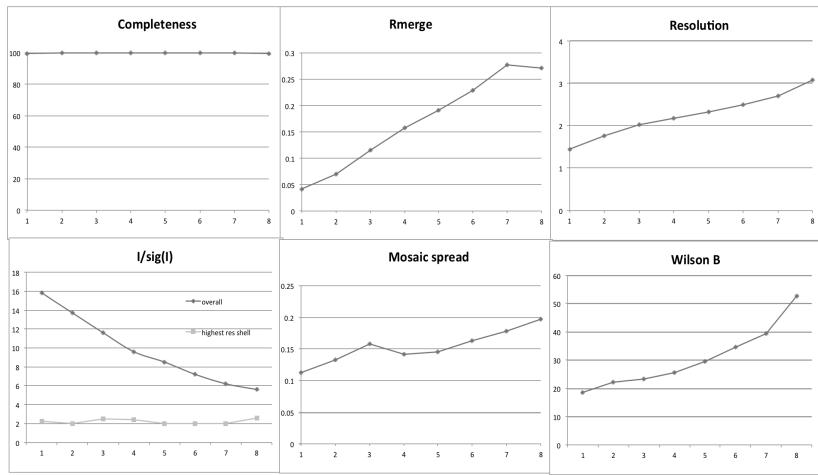
As the engineered cavity is occupied by the wildtype tryptophan W191 in this case, a single site fit was used to determine a K_d of 25.6 mM for the secondary site. The R^2 for this single-site binding hyperbola is 0.9908.



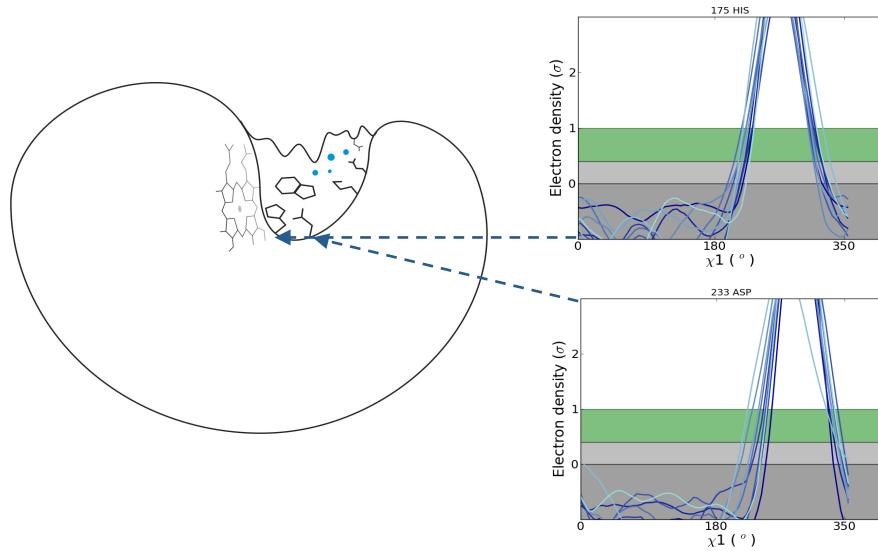
SI Figure 7 CONTACT network analysis¹⁴ shows no steric coupling of the cryptic site and the heme (both shown in red sticks).

SI Figure 8

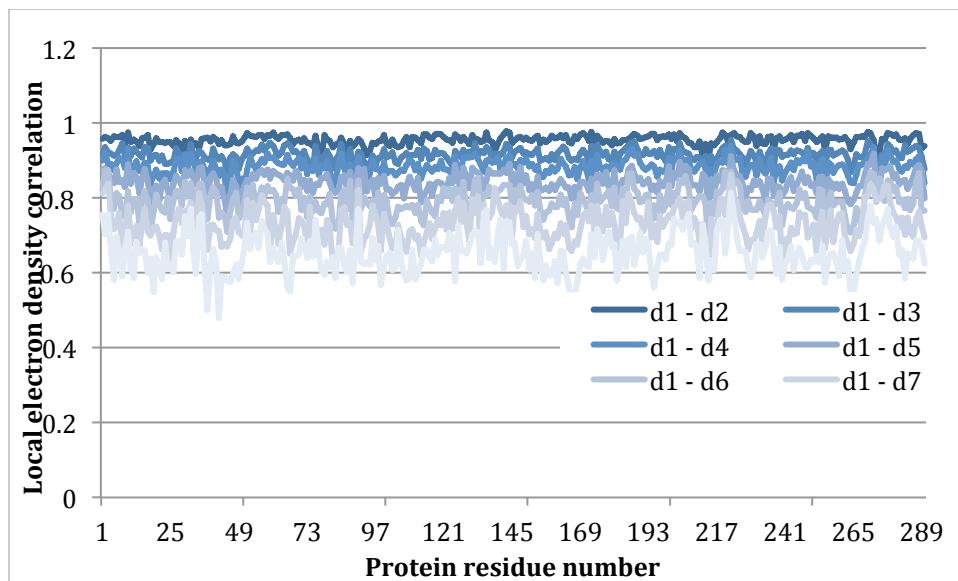
A)



B)



C)



SI Figure 8 A) Data collection statistics across 8 consecutively collected datasets. B) Ringer plots show no significant change in side-chain conformations of residues near the heme-Fe in the course data collection with increasing doses of radiation. C) Local electron density correlation (LDC) shows a negligible impact of radiation damage on local structure at RT between consecutive datasets. Correlation coefficients between the electron densities of 8 datasets continuously collected on a single crystal at room temperature, compared to the first dataset collected, enumerated over all residues. Although the initial unrefined map was used the correlation is very good excluding radiation damage as a source for the differences between the RT-cryo pairs.

Supplemental References

1. Rocklin, G.J. et al. Blind prediction of charged ligand binding affinities in a model binding site. *J Mol Biol* (2013).
2. Winter, G., Loble, C.M. & Prince, S.M. Decision making in xia2. *Acta crystallographica. Section D, Biological crystallography* **69**, 1260-73 (2013).
3. McCoy, A.J. et al. Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007).
4. Murshudov, G.N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D* **67**, 355-367 (2011).
5. Afonine, P.V. et al. Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallographica Section D* **66**, 1153-1163 (2010).
6. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D* **60**, 2126-2132 (2004).
7. Fraser, J.S. et al. Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16247-52 (2011).
8. Frauenfelder, H. et al. Thermal expansion of a protein. *Biochemistry* **26**, 254-61 (1987).
9. Frauenfelder, H., Petsko, G.A. & Tsernoglou, D. Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* **280**, 558-63 (1979).
10. Lang, P.T. et al. Automated electron-density sampling reveals widespread conformational polymorphism in proteins. *Protein science : a publication of the Protein Society* **19**, 1420-31 (2010).
11. van den Bedem, H., Bhabha, G., Yang, K., Wright, P.E. & Fraser, J.S. Automated identification of functional dynamic contact networks from X-ray crystallography. *Nature methods* **10**, 896-902 (2013).
12. Chelvanayagam, G., Knecht, L., Jenny, T., Benner, S.A. & Gonnet, G.H. A combinatorial distance-constraint approach to predicting protein tertiary models from known secondary structure. *Folding & design* **3**, 149-60 (1998).
13. Barelier, S. et al. Roles for ordered and bulk solvent in ligand recognition and docking in two related cavities. *PLOS ONE* (2013).
14. van den Bedem, H., Bhabha, G., Yang, K., Wright, P.E. & Fraser, J.S. Automated identification of functional dynamic contact networks from X-ray crystallography. *Nat Methods* **10**, 896-902 (2013).