

Accessing protein conformational ensembles using room-temperature X-ray crystallography

James S. Fraser^{a,1}, Henry van den Bedem^b, Avi J. Samelson^a, P. Therese Lang^a, James M. Holton^{c,d}, Nathaniel Echols^d, and Tom Alber^{a,2}

^aDepartment of Molecular and Cell Biology, University of California, Berkeley, CA 94720-3220; ^bJoint Center for Structural Genomics, Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025; ^cDepartment of Biochemistry and Biophysics, University of California, San Francisco, CA 94158-2330; and ^dLawrence Berkeley National Laboratory, 1 Cyclotron Road, Building 64R0121, Berkeley, CA 94720

Edited by Gregory A. Petsko, Brandeis University, Waltham, MA, and approved July 27, 2011 (received for review July 13, 2011)

Modern protein crystal structures are based nearly exclusively on X-ray data collected at cryogenic temperatures (generally 100 K). The cooling process is thought to introduce little bias in the functional interpretation of structural results, because cryogenic temperatures minimally perturb the overall protein backbone fold. In contrast, here we show that flash cooling biases previously hidden structural ensembles in protein crystals. By analyzing available data for 30 different proteins using new computational tools for electron-density sampling, model refinement, and molecular packing analysis, we found that crystal cryocooling remodels the conformational distributions of more than 35% of side chains and eliminates packing defects necessary for functional motions. In the signaling switch protein, H-Ras, an allosteric network consistent with fluctuations detected in solution by NMR was uncovered in the room-temperature, but not the cryogenic, electron-density maps. These results expose a bias in structural databases toward smaller, overpacked, and unrealistically unique models. Monitoring room-temperature conformational ensembles by X-ray crystallography can reveal motions crucial for catalysis, ligand binding, and allosteric regulation.

protein conformational dynamics | energy landscape | Ringer | qFit

Macromolecular X-ray crystallographic diffraction experiments provide powerful insights into the relationship between structure and biological function. Although macromolecules populate vast ensembles of alternative conformational substates (1), crystallographic models depicting the major average conformation have provided foundational ideas about the mechanisms of biochemical reactions. By slowing radiation damage to the sample, crystal cooling has catalyzed a revolution in structural biology, enabling structure determinations from tiny crystals using bright synchrotron X-ray sources (2–4). It is estimated that more than 95% of the >65,000 crystal structures deposited in the Protein Data Bank (PDB) are based on cryogenic data (5).

Crystal cooling is generally thought to introduce little bias in the functional interpretation of structural results. Some investigators have suggested that the standard practice of plunging crystals into liquid nitrogen and collecting X-ray diffraction data at 100 K traps a representative set of conformations populated at room temperature (6, 7). In contrast, early structural comparisons by Petsko, Frauenfelder, and colleagues indicated that cooling myoglobin crystals causes a small reduction in the protein volume due to anisotropic displacements of atomic positions and subtle changes of contacts between α -helices (8). A landmark study of crystalline ribonuclease A (RNaseA) by Tilton, Petsko, and coworkers revealed diverse temperature-dependent changes in the average structure, ranging from shrinkage at low temperatures to increased loop disorder at 47 °C (9). A recent analysis comparing 15 room-temperature and cryogenic crystal structures documented that cryocooling generally increases lattice contacts and reduces protein volumes but causes only small changes in crystallographic models (10).

In contrast to these small conformational perturbations, the dynamic characteristics of proteins in crystals are exquisitely sensitive to temperature (11, 12). After accounting for static structural differences between molecules, model errors, and crystal lattice defects in myoglobin crystals, Frauenfelder and Petsko demonstrated widespread temperature-dependent decreases in the harmonic atomic vibrations measured by B-factors. Based on the distribution of these differences, they proposed that minor conformational substates with alternative side-chain conformations are involved in transient ligand-entry pathways (13). In crystalline RNaseA in liquid solvent, a “glass transition” in the protein between 212 and 228 K reduces B-factors, increases the quality of core packing, and eliminates motions needed for inhibitor binding to the active site (9, 14). The cooling-induced reductions in B-factors suggest that cryogenic structures adopt less variable conformations.

Indirect measures such as B-factors or protein-volume changes, however, do not distinguish whether cryocooling simply reduces vibrational motions or biases the overall conformational ensemble. A theoretical analysis by Halle suggests that the crystal cryocooling process, which typically takes up to a second, is too slow to trap the room-temperature equilibrium distribution of protein and solvent configurations (15). According to Halle’s dynamic quenching theory, changes in the energy landscape upon cooling can cause specific changes in the relative populations of alternative conformations that are not apparent from B-factor analysis or the standard refinement of unique models. Consistent with this view, we found previously that catalytically essential, alternative conformations of the proline isomerase, CypA, detected in solution using NMR are populated in crystals at room temperature but absent at cryogenic temperature (16). This temperature-dependent difference suggested that functionally important mobile regions of proteins might be particularly susceptible to annealing into a unique conformation during the cooling process.

To test the idea that crystal cryocooling changes protein ensembles, we examined 30 proteins for which high-resolution X-ray data were collected previously at ambient and cryogenic

Author contributions: J.S.F. and T.A. designed research; J.S.F., H.v.d.B., and A.J.S. performed research; J.S.F., H.v.d.B., P.T.L., J.M.H., and N.E. contributed new reagents/analytic tools; J.S.F. and H.v.d.B. analyzed data; and J.S.F., H.v.d.B., A.J.S., P.T.L., J.M.H., N.E., and T.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The structure factors have been deposited in the Protein Data Bank, www.pdb.org.

¹Present address: California Institute of Quantitative Biosciences (QB3) and Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94158

²To whom correspondence should be addressed. E-mail: tom@ucxray.berkeley.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1111325108/-DCSupplemental.

temperatures. To access the conformational ensembles, we computationally sampled the electron-density maps using the programs qFit, to build and refine multiconformer models (17), and Ringer, to define unmodeled alternative side-chain conformations (18). These methods sample values of electron density that are usually ignored during model building to discover alternative conformations and distinguish these substates from harmonic motions of the average structure. Our results show that crystal cryocooling systematically distorts the repertoire of accessible conformations. Conformational ensembles based on room-temperature data generally provide a fuller description of the physiological macromolecular mechanisms than traditional models based on cryogenic data.

Results

Automated Multiconformer Refinement Enables Temperature Comparisons. To compare conformational ensembles, we identified 30 proteins with available high-resolution (2.0 Å or better) X-ray diffraction data collected at ambient and cryogenic temperatures (Table S1). For 17 of the 30 pairs (including our data on CypA), the same investigators collected the X-ray datasets. Additionally, 20 of the 30 pairs were crystallized in nearly identical conditions (differing only slightly in concentrations of buffer, precipitant, salt, and cryoprotectant). The datasets have R_{merge} values indicating no unusual radiation damage (Fig. S1), and we excluded data that were collected specifically to examine the effect of temperature on radiation damage.

To reduce biases introduced by different refinement programs and model builders, we rebuilt all 60 models using qFit (17), which automates the discovery and building of alternative polypeptide conformations, allowing for coupled side-chain-backbone flexibility. The multiconformer qFit models were further refined with phenix.refine (19) to ensure that all models experienced similar restraints on geometry, B-factors and solvent, improving R_{free} values by an average of 1.0% relative to the deposited structures. Because the mean resolution is high for both the room-temperature (1.44 Å) and cryogenic-temperature (1.30 Å) datasets, the estimated coordinate errors are low (0.05–0.22 Å with a mean of 0.13 Å). In all cases, the coordinate error is smaller than the rmsd for equivalent non-hydrogen backbone atoms (mean = 0.34 Å), suggesting that small but significant temperature-dependent differences exist between the models.

To further access conformational ensembles, we used the program Ringer (18). Ringer samples electron density around side-chain dihedral angles and identifies peaks representing discrete alternative conformations at levels of electron density traditionally considered to be dominated by noise. This information can provide an estimate of the distribution and relative populations of numerous unmodeled alternative side-chain conformations. The conformations identified using a combination of Ringer and manual model building agree generally with those identified using qFit (Fig. 1 and Fig. S1B).

Cryocooling Remodels Side-Chain Conformational Distributions. To test the idea that cryocooling affects more than the harmonic vibrations (Fig. S24) of the conformational ensemble, we used Ringer to sample the electron density around all residues with a rotatable χ_1 angle (4,008 residues from 30 proteins). For each residue, we calculated a χ_1 -correlation coefficient between the room-temperature and cryogenic electron-density distributions. The χ_1 -correlation coefficient establishes the conformational basis for differences in side-chain mobility with minimal dependence on map scale or resolution. If conformations are unperturbed and residues only experience a loss of thermal motion, then the electron-density distributions of the room-temperature and cryogenic datasets are highly correlated. At room temperature, for example, Asn106 in asparaginase has a lower electron density peak and wider edges that reflect increased thermal mo-

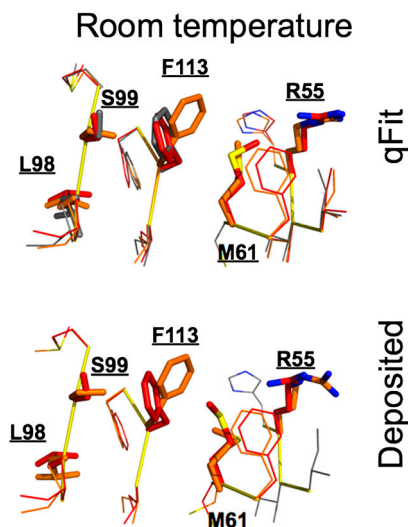


Fig. 1. qFit and Ringer automate the discovery of alternative conformations. Automated qFit (top) and manual Ringer-guided (bottom) model building for CypA into room-temperature electron density reveals the temperature dependence of interconverting side-chain conformations linked to catalysis for CypA. Both procedures identify alternative conformations (colored in red, orange, and gray) for dynamic network residues using the room-temperature X-ray data (residues Leu98, Ser99, Phe113, Met61, Arg55). For some residues (for example Phe60), qFit builds alternative conformations to optimize geometry within a single rotameric substate.

tion, which is well modeled by B-factors and yields Ringer plots with a correlation coefficient of 0.99 (Fig. 24i). Most residues (69.3%) have χ_1 correlation coefficients greater than 0.9, which indicates the primary effect of cryocooling is to reduce thermal motion around conformations with similar populations at both temperatures (Fig. 2B).

However, Ringer plots with lower correlation coefficients highlight a variety of complex, cooling-induced changes to side-chain conformational distributions (Fig. 24). For example, Leu98 from CypA (Fig. 24ii) and Ser109 from superoxide reductase (Fig. 24iii) demonstrate alternative conformations that have reduced populations at cryogenic temperatures. Additional residues with Ringer correlation coefficients below 0 indicate that cryocooling induces a new conformation that was not significantly occupied at room temperature. For example, trypsin Val53 in the complex with bovine pancreatic trypsin inhibitor (BPTI) switches from the 180° rotamer to the previously unoccupied –60° rotamer (Fig. 24iv). Using a conservative correlation cutoff of 0.85 to determine if cryocooling changes the side-chain conformational distribution (by reducing the populations of alternative conformations or switching rotamers), we found that 18.9% of all residues are remodeled at χ_1 by cryogenic data collection (Fig. 2B). In agreement with the idea that solvent interactions are altered at cryogenic temperatures (9, 14), residues at structurally unconstrained solvent-accessible sites are most susceptible to these effects. However, cryocooling remodels the χ_1 angle for 7.8% of all buried residues (Fig. 2B). Because differences in the models precluded simple comparisons of the Ringer correlations at other side-chain dihedral angles, these results represent a lower limit on the effect of cryocooling on side-chain conformations that can be detected in the electron density.

To assess the effects of cryocooling at higher χ -angles and account for main chain variations, we examined alternative conformations from the qFit models. For each residue at each χ -angle, we calculated the maximum absolute change in rotamer occupancy between the room-temperature and cryogenic ensembles. Importantly, this measure can compare any combination of multiconformer or single-conformer residues. We found that 13.1% of residues have at least a 20% change in occupancy at χ_1

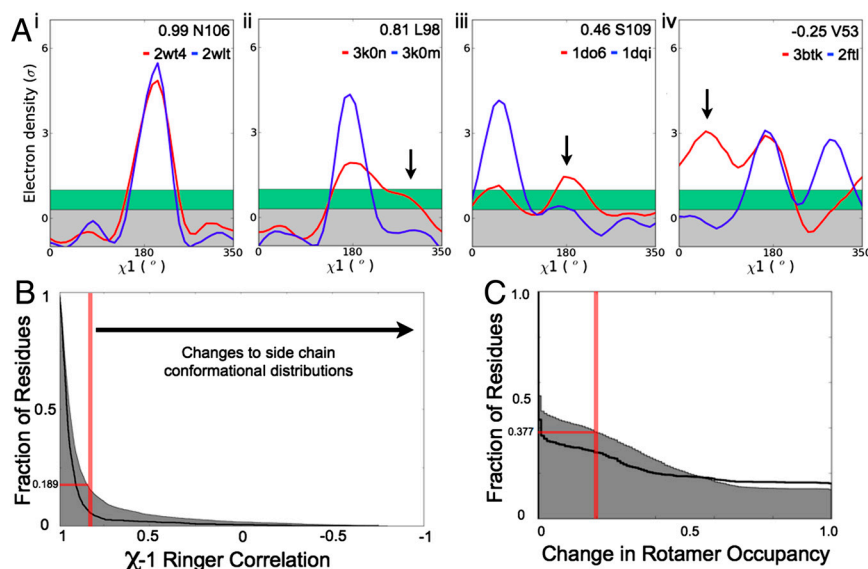


Fig. 2. Cryocooling changes vibrational amplitudes and conformational distributions. (A) Ringer plots and correlations of electron-density distributions reflect both vibrational and conformational changes induced by crystal cryocooling. Correlation coefficients of Ringer plots based on the room-temperature density (red) and the cryogenic density (blue) in cross-validated maximum-likelihood-weighted, $2mF_o - DF_c$ maps classify changes. Electron-density levels between $0.3-1\sigma$ (shaded green) are enriched in alternative conformations compared to lower electron-density levels ($<0.3\sigma$, shaded gray). (i) Asn106 of asparaginase has a high (0.99) correlation coefficient with reduced thermal motion and no changes to alternative conformations. Intermediate correlations for (ii) Leu98 of CypA and (iii) Ser109 of superoxide reductase reflect differences in the relative populations of alternative conformations (arrows). (iv) The characteristic double peaks for Val53 C γ 1 and C γ 2 of trypsin in the complex with BPTI shift upon cryocooling. The negative correlation coefficient reflects the occupancy of a new rotamer. (B) A cumulative histogram of Ringer correlations for all residues (shaded gray area) and buried residues (black line) reveal that cryocooling alters the distribution of conformational substates. Of all residues, 18.9% have correlations coefficients less than 0.85 (red vertical line), indicating significant changes to χ_1 electron-density distribution. (C) A cumulative histogram of the maximum absolute change in rotamer occupancy from qFit models for all residues (shaded gray area) and buried residues (black line) reveal that 37.7% of residues have a 20% or greater change in refined occupancy (red vertical line).

(Fig. S2B). As we sample further away from the backbone, the percentage of residues that are remodeled by cryocooling increases from 24.5% at χ_2 to 41.9% at χ_3 (Fig. S2B). Consistent with the Ringer analysis, the effect of cryocooling is greatest at solvent-exposed sites. Buried residues, however, are also altered by the cooling process and display a similar trend of increasing variability from 4.7% at χ_1 to 16.6% at χ_2 to 35.9% at χ_3 as the distance from the backbone increases. In total, we observed at least a 20% occupancy shift in rotamer for 37.7% of all residues and 29.0% of buried residues (Fig. S2C).

For each residue, we also computed the maximum rmsd between alternative side-chain conformations starting at the C β atom. The room-temperature models have significantly greater average maximum rmsds (0.75 Å all residues, 0.49 Å buried residues, 0.84 Å solvent-exposed residues) than the cryogenic models (0.64 Å all residues, 0.39 Å buried residues, 0.72 Å solvent-exposed residues). These results support the idea that cryocooling remodels the conformational ensembles of both solvent-exposed and buried side chains.

Solvent and Lattice Compression Remodels Surface Residues. To examine whether the driving forces that remodel side-chain ensembles are consistent with previous studies of the temperature dependence of crystal structures (10, 20), we measured changes of unit-cell volume, protein volume, and lattice contacts. In our dataset, 27 of 30 crystal unit-cell volumes contract upon cryocooling (mean contraction = 3.4%), and the unit-cell volumes generally shrink more than the protein volume (mean contraction = 1.4%) (Fig. 3A). Although most atoms move toward the center of mass upon cryocooling, isotropic compression alone is insufficient to account for the coordinate differences, as previously observed for myoglobin (8) and alpha-lytic protease (6). Increases in protein volume can occur by the formation of new lattice contacts (Fig. S3), resulting in expulsion or thermal contraction of solvent (20). To assess the effects on lattice contacts, we measured occupancy-weighted non-hydrogen atomic contacts at a 3.5-Å distance

cutoff (Fig. S4A). Increases in lattice contacts (mean = +17%) suggest that cryocooling creates a larger contact area by reducing the solvent channels. Thus, cryocooling increases the crystalline nature of the protein, changing the relatively loose association between molecules at room temperature into a more stable lattice that buries more surface area. This lattice stabilization, combined with reduced thermal motion at lower temperatures, might explain improvements in diffraction resolution observed at cryogenic temperatures. Alternatively, cryocooling in unsuitable solvents can cause dramatic losses in diffraction quality if

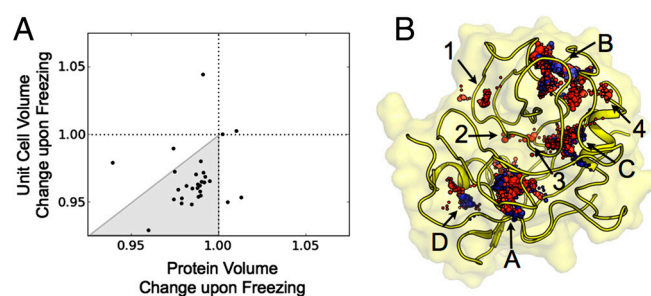


Fig. 3. Crystal cooling contracts unit cells and proteins and increases packing quality. (A) The distribution of changes in unit-cell volumes and protein volumes reveals asymmetric alterations upon cryocooling. Protein volume decreases (cryogenic/room-temperature ratio <1.0) are generally smaller than the changes in unit-cell volumes (gray shaded area). This suggests that unit-cell volume decreases (cryogenic/room-temperature ratio <1.0) result from expulsion of solvent from channels in addition to protein compression. Expansions are created by cooling-induced crystal contacts or subdomain reorientations (Fig. S2). (B) Visualizing room-temperature (red) and cryogenic (blue) packing defects for CypA indicates common defects (indicated by letters) and defects unique to the room-temperature model (indicated by numbers). Defect A is present at both temperatures and is associated with a dynamic loop important for HIV-Capsid binding selectivity. Defect 1 is associated with an alternative conformation of Phe113. Defects 2 and 3 are associated with functional alternative conformations of Leu 98 and Ser99.

the lattice cannot rearrange by forming new contacts as solvent is expelled.

Quantitative Improvements in Protein Packing Remodels Buried Residues. Changes to crystal contacts and surface mobile loops, however, do not explain why buried residues change in response to crystal cryocooling. Because cryocooling decreases the protein volume without changing the number of atoms, we reasoned that a major effect would be to change the quality of packing. To test this hypothesis, we used RosettaHoles2 (21), which calculates the distribution of small interior packing defects and large voids. Our results confirmed that the cryogenic structures have superior packing compared to the isomorphous high-resolution room-temperature structures (Fig. S4B). The rubredoxin models from joint neutron and X-ray studies, where the incomplete deuteration of the cryogenic structure may be incorrectly modeled during refinement (22), were the only exception to this trend. A possible caveat to this pattern is that the cryogenic structures are generally determined at higher resolution, and high-resolution refinement can produce a better description of packing. However, even for the five structures where the room-temperature data were collected to higher resolution, the cryogenic structures demonstrate superior packing.

To examine how this loss of packing defects relates to function, we examined the spatial distribution of temperature-dependent packing defects in CypA. Several packing defects (indicated by letters in Fig. 3B) occur in both the cryogenic and room-temperature models, which may enable functional protein flexibility. For example, Defect A, which is present at both temperatures, abuts a loop that participates in a conformational exchange process observed by NMR (23). Mutations that change the local packing surrounding this void impact the loop conformation and modify the affinity of primate CypA homologs for the HIV capsid (24). In addition to the cavities present at both temperatures, the room-temperature structure contains several small packing defects (indicated by numbers in Fig. 3B) that are absent at cryogenic temperatures due to the compression of the protein. Three temperature-dependent small cavities occur adjacent to residues with alternative conformations that are essential for catalysis (16) (Fig. S5).

The Room-Temperature Ensemble Reveals Fluctuations in H-Ras. In CypA, crystal cryocooling preferentially biases alternative side-chain conformations responsible for conformational exchange detected by NMR relaxation experiments (16). A coordinated analysis of X-ray and NMR data uncovered a correspondence between the regions of the protein that are sensitive to cooling

and those that are mobile in solution. To test the generality of this idea, we analyzed the small GTPase H-Ras bound to the GTP-analog, GMPPNP, by comparing high-resolution room-temperature (1.31 Å) and cryogenic (1.26 Å) X-ray data (25) to NMR relaxation dispersion experiments (26).

Experimental and computational studies (27, 28) have established the basis for long-range conformational coupling and the functions of effector proteins in the GDP- to GTP-bound allosteric switch in small GTPases such as Ras. NMR studies of H-Ras (26) suggested an additional intrinsic conformational rearrangement within the GTP-bound state (Fig. 4A). However, the high-resolution crystal structure of H-Ras bound to GMPPNP paradoxically revealed a single conformation, which is not poised for hydrolysis. This catalytically incompetent conformation left unresolved the basis for the fluctuations observed by NMR. Because this conformational exchange process is not structurally defined, it is unclear if these motions promote interactions such as binding to effectors such as a GTPase Activating Protein (GAP) (29) or populate catalytically competent conformations.

In the ensembles of H-Ras:GMPPNP, the catalytically incompetent conformation dominates at both cryogenic and room temperatures. Unexpectedly at room temperature, we detected an alternative, catalytically competent Gln61 conformation at low values of electron density below the 1σ level traditionally used for model building (Fig. 4B and C). In the minor conformation of the active site, Gln61 switches rotamers and positions a water molecule that is poised to attack the terminal phosphate (Fig. 4C). This new Gln61 rotamer resembles the active conformation stabilized by the Ras-GAP complex with the transition state analog, GDP- AlF_3 (29, 30) (Fig. 4B and Fig. S6A).

In agreement with the NMR relaxation data (26), many residues surrounding Gln61 and extending to helix 3 show temperature-induced changes in side-chain conformational distributions (Fig. 4C and Fig. S6B). Temperature-dependent differences in these areas are apparent in electron-density maps calculated in different ways and are not likely due to changes in radiation damage or global crystal quality (Fig. S7). One network identified by NMR is centered on His94 in helix 3 (Fig. 4A). Strikingly, the room-temperature electron-density map revealed at low levels of electron density alternative conformations in residues along helix 3—including two His94 conformations—whereas the cryogenic map contains only a single conformation (Fig. 5A). The minor substates observed at room temperature encompass a subtle repositioning of helix 3 associated with the new rotamer distributions of His94 and neighboring residues.

To understand how H-Ras moves from a diverse ensemble at room temperature to a more unique structure upon cooling, we

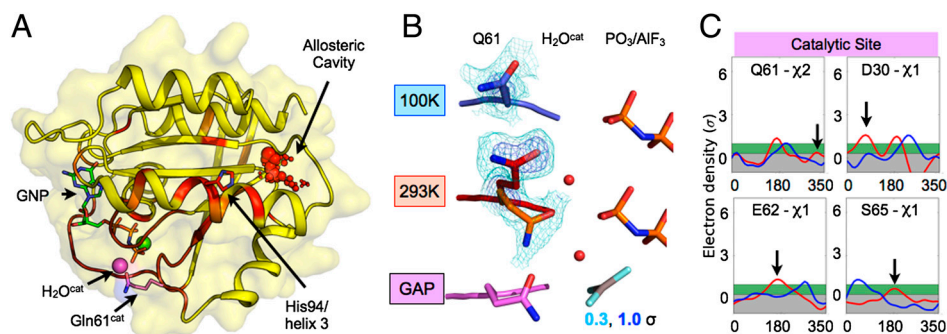


Fig. 4. Cryocooling quenches an intrinsically flexible network in H-Ras. (A) Residues with detectable conformational exchange by NMR (26) (brown, orange, and red) mapped onto the structure of GMPPNP-bound H-Ras (PDB ID code 5P21). Residues colored red were fit to a collective exchange rate, and residues colored brown were too broadened to permit classification into a collective exchange process. Gln61 is thought to position a catalytic water molecule ($\text{H}_2\text{O}^{\text{cat}}$) during intrinsic hydrolysis of GTP. (B) Models built into room temperature (red and orange), but not the cryogenic (blue), electron-density maps reveal two conformations of Gln61. Gln61 fluctuates to a catalytically competent conformation resembling the GDP- AlF_3 transition state complex stabilized by GAP binding (magenta). Electron-density maps are shown at 1.0 (blue) and 0.3 (cyan) σ contours. The X-ray datasets are derived from a single study of H-Ras crystals grown in a single condition. (C) Ringer electron-density sampling reveals Gln61 and surrounding residues in the active site display conformational differences between the room-temperature structure (red) and cryogenic structure (blue).

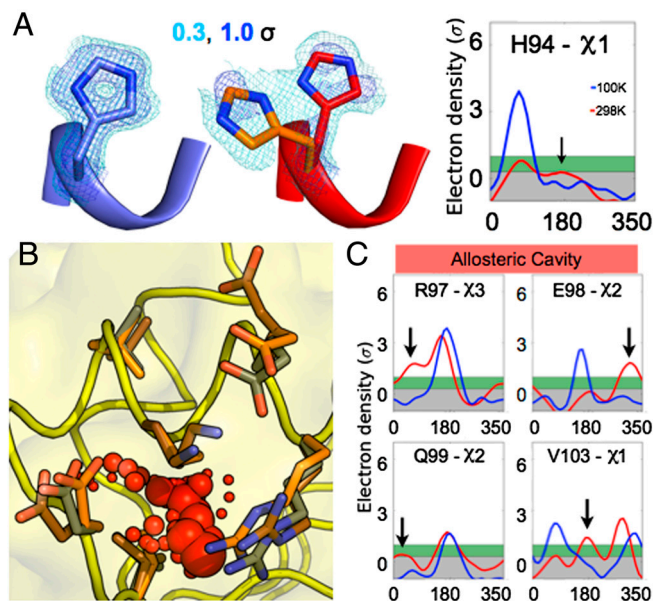


Fig. 5. Multiple conformations in the active site and allosteric cavity of H-Ras. (A) Room-temperature (red and orange), but not the cryogenic (blue), electron-density maps reveal two conformations of His94. Rotameric switching of His94 may influence the chemical shifts of surrounding residues, explaining how NMR relaxation dispersion signals across helix 3 can arise with minimal changes in backbone geometry. (B) An allosteric cavity (red spheres) at the C-terminal end of helix 3 is present only at room temperature. The cavity is transiently filled by alternative conformations (primary conformation in brown, alternative conformations in orange and gray) that become the major conformation and fill the cavity during crystal cryocooling. (C) Ringer electron-density sampling reveals temperature-dependent, side-chain conformational differences at the allosteric cavity, including alternative conformations of Arg97 and Lys101.

examined changes in packing. At the end of helix 3, a cavity collapses upon cryocooling, which causes changes in relative populations of alternative conformations of the surrounding residues (Fig. 5B and C). The movement of helix-3 residues Arg97, Glu98, and Lys101 into the cavity at cryogenic temperatures causes the small backbone rearrangements that stabilize the single rotamer of His94. Collapse of the cavity propagates through a local backbone shift, which reduces side-chain disorder. In agreement with the NMR results, this side-chain disorder traverses the protein through residues that are flexible in solution to the active site 20 Å away (Fig. 4A and Fig. S6B). Thus, unlike the jigsaw-puzzle-like packing of the cryogenic structure, the room-temperature ensemble reveals alternative conformations across a network that corresponds structurally with the conformational exchange processes observed by NMR.

Discussion

By defining previously undetected protein conformational ensembles rather than unique structural models, our analysis suggests that the nearly universal practice of cryocrystallography shifts the intrinsic populations of conformers. Pervasive and specific modifications to protein side-chain distributions occur upon crystal cryocooling. These biases across 30 pairs of structures are associated with compression of the protein, increases in lattice contacts, and reduction of thermal motion. For CypA and H-Ras, crystal cooling restricts coupled motions precisely in flexible regions detected in solution using NMR. These patterns are unlikely to arise from differences in radiation damage between room temperature and cryogenic data collection.

Because alternative conformations reflect nearly iso-energetic substrates (31), changes in the energy landscape induced by cooling can readily shift populations in the ensemble. Low-enthalpy, low-entropy conditions at low temperatures change solvent inter-

actions, reduce packing defects, and redistribute alternative conformations. Our results support the theoretical prediction that crystal cooling is too slow to trap the equilibrium side-chain distributions at both solvent-exposed and buried positions of proteins (15). Consistent with Halle's dynamic quenching theory (15), cryocrystallography causes anisotropic and idiosyncratic effects on the contraction of each protein and the quality of local packing (Fig. 3). These tendencies, including the tightness of core packing and reduced protein volume, may percolate throughout computational biophysics, where accuracy is often judged by recapitulating features of high-resolution cryogenic X-ray structures.

Understanding how individual proteins and solvents respond to crystal cryocooling may help reconcile the static images derived from traditional X-ray crystallography experiments with the dynamic view of proteins from solution NMR experiments (32) or comparisons of multiple independent crystal structures (33). X-ray crystallography is intrinsically an ensemble measurement, even though unique models are normally used to communicate the results of X-ray diffraction experiments. By detecting minor populations at electron-density levels often thought to be dominated by noise, recently developed methods including qFit and Ringer enhance access to alternative conformations from individual crystals. Ensembles detected in room-temperature electron-density maps of CypA (16) and H-Ras (Figs. 4 and 5) structurally define important, functional, conformational exchange processes that also occur in solution.

Our analysis suggests that crystal cryocooling restricts H-Ras from sampling conformations that could promote GTP hydrolysis. Electron-density maps calculated with data obtained at room temperature reveal hidden alternative conformations in which Gln61 populates the catalytically competent conformation. The discordance between the rate of NMR fluctuations (microsecond-millisecond) and the rate of turnover (minutes-hours) suggests that this conformational transition does not limit the rate of intrinsic GTP hydrolysis. The presence of multiple independently correlated processes, indicated by NMR and our analysis of the X-ray data, implies that the inefficiency of intrinsic hydrolysis may result from uncorrelated motions around the catalytic center. Alternatively, the conformational plasticity of Gln61 may facilitate a conformational selection mechanism of GAP binding to the catalytically competent conformations. This model makes the prediction that perturbations along the network, such as binding of effectors or mutations along helix 3, could increase conformational coupling to alter the functional properties of H-Ras. Indeed, adventitious Ca^{2+} and acetate binding in the allosteric cavity stabilizes the active conformation of the Gln61 network in crystals (34) (Fig. S6C), and a mutation that adds hydrophobic bulk to this cavity (Lys101Leu) switches cellular morphologies associated with Ras activation of the Raf kinase (35). Raf interacts with H-Ras near the active site over 20 Å away from the site of the mutation, functionally linking an expanded cavity to the conformation of the active site (Fig. S6C). ^{31}P NMR measurements indicate that mutations that disrupt hydrogen bonds along this network alter interactions between H-Ras and the GTP-analogue (36). In addition, mutations of H-Ras paralogs that abut or are located on helix 3 (Arg68Thr, Asp69Asn, Met72Ile, Asp92Asn, Asp92Tyr, Arg97Gly) have been identified in cancers (37), supporting the idea that the network is critical for the function of H-Ras (Fig. S6C).

Our results suggest that X-ray crystallography—along with simulations (38), sequence analyses (39), and NMR experiments (40)—can offer increasing access to conformational ensembles required for protein functions. Ringer and qFit are best employed using X-ray data to at least 2.0-Å resolution, and almost half of recent PDB depositions meet this threshold. In the cases we have examined, linking ensemble analysis to the solution distribution of substrates requires room-temperature X-ray data. However, in recent years, room-temperature X-ray data collection has

become quite rare, and cryogenic structures dominate the PDB. In our experience, many crystals that diffract to high-resolution at cryogenic temperatures can be used to collect comparable room-temperature data at a synchrotron. Our results suggest that effort should be put toward the redeployment of data collection techniques at room temperature with careful attention to standardizing practices that mitigate radiation damage. A combined strategy of ensemble analysis and room-temperature X-ray data collection can be applied to many proteins to define conformational substates linked to ligand binding, catalysis, and allosteric regulation.

Materials and Methods

Structure factors in cif format were downloaded from the PDB and converted to mtz format with phenix.cif_as_mtz (19). When no R_{free} set was deposited or could be extracted, we chose a test set using the standard parameters in phenix.refine (19). Structures were rebuilt using qFit (17) and then refined for five additional cycles in phenix.refine, with occupancy refinement and automated solvent picking. For all models, we included riding protein hydrogen atoms in refinement. Datasets at <1.6 -Å resolution were refined with anisotropic B-factors for all nonsolvent and nonhydrogen atoms. Unit-cell volumes were calculated with matthews_coef from CCP4 (41). Protein volumes, calculated with 3V (42), and packing scores, calculated with RosettaHoles2 (21), are based on all matching nonhydrogen atoms of the primary conformation. Residues with side-chain-up half-sphere exposure (radius = 13 Å) greater than 24 were considered buried (43). All structural figures were prepared with PyMol.

Ringer analyses of χ^2 electron density were performed by sampling from the major conformation in unfilled cross-validated, maximum-likelihood-

weighted, $2mF_o - DF_c$ maps created by phenix.maps with a grid set to 1/5 of resolution. Although the maps are from isomorphous crystals at high resolution, to reduce the effect of map scaling and resolution, we used the correlation coefficients of Ringer electron density distributions in 10° increments to determine the behaviors of side-chain conformational distributions. Multiconformer residues from qFit models were analyzed by summing the occupancy of each alternative conformation into the nearest rotameric bin (60° , 180° , 300°), for individual χ -angle analysis or into the nearest rotamer (44) assigned by phenix.rotalyze, for the analysis across all angles. The maximum absolute differences in rotameric occupancy between the room-temperature and cryogenic ensembles range from 0, indicating that the conformational distribution is identical, to 1, indicating that there is no agreement in rotamer populations. The average maximum rmsd between room temperature and cryogenic models increased by 0.11 Å ($p = 1.7 \times 10^{-17}$) for all residues, 0.09 Å ($p = 3.6 \times 10^{-7}$) for buried residues, and 0.12 Å ($p = 3.0 \times 10^{-12}$) for solvent-exposed residues. A two-sided paired t test was used to determine significance.

ACKNOWLEDGMENTS. We thank Dale Tronrud, Brian Matthews, and Axel Scheidig for providing structure factors for room-temperature datasets; Will Sheffler and David Baker for providing code and guidance for RosettaHoles2; Mickey Kosloff, Dorothee Kern, Judith Klinman, Tanja Kortemme, Irwin Kuntz, Susan Marqusee, and Dan Tawfik for stimulating discussions and helpful comments. J.S.F. was supported by fellowships from the Natural Sciences and Engineering Research Council and the National Science Foundation; H.v.d.B. was supported by the National Institute of General Medical Sciences Protein Structure Initiative (U54GM094586) at the Joint Center for Structural Genomics; J.M.H. was supported by grants from the National Institutes of Health (NIH) (U54GM094625 and P50GM073210); this work was supported by a grant from the NIH (R01GM48958) to T.A.

- Karplus M, Kuriyan J (2005) Molecular dynamics and protein function. *Proc Natl Acad Sci USA* 102:6679–6685.
- Southworth-Davies RJ, Medina MA, Carmichael I, Garman EF (2007) Observation of decreased radiation damage at higher dose rates in room temperature protein crystallography. *Structure* 15:1531–1541.
- Low BW, Chen CC, Berger JE, Singman L, Pletcher JF (1966) Studies of insulin crystals at low temperatures: Effects on mosaic character and radiation sensitivity. *Proc Natl Acad Sci USA* 56:1746–1750.
- Hope H, et al. (1989) Cryocrystallography of ribosomal particles. *Acta Crystallogr B* 45:190–199.
- Garman E (2003) 'Cool' crystals: Macromolecular cryocrystallography and radiation damage. *Curr Opin Struct Biol* 13:545–551.
- Rader SD, Agard DA (1997) Conformational substates in enzyme mechanism: The 120 K structure of alpha-lytic protease at 1.5 Å resolution. *Protein Sci* 6:1375–1386.
- Coureaux PD, Fan ZP, Stojanoff V, Genick UK (2008) Picometer-scale conformational heterogeneity separates functional from nonfunctional states of a photoreceptor protein. *Structure* 16:863–872.
- Frauenfelder H, et al. (1987) Thermal expansion of a protein. *Biochemistry* 26:254–261.
- Tilton RF, Jr, Dewan JC, Petsko GA (1992) Effects of temperature on protein structure and dynamics: X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320 K. *Biochemistry* 31:2469–2481.
- Juergens DH, Matthews BW (2001) Reversible lattice repacking illustrates the temperature dependence of macromolecular interactions. *J Mol Biol* 311:851–862.
- Weik M, Colletier JP (2010) Temperature-dependent macromolecular X-ray crystallography. *Acta Crystallogr D Biol Crystallogr* 66:437–446.
- Fraser JS, Jackson CJ (2010) Mining electron density for functionally relevant protein polyserism in crystal structures. *Cell Mol Life Sci* 68:1829–1841.
- Frauenfelder H, Petsko GA, Tsernoglou D (1979) Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* 280:558–563.
- Rasmussen BF, Stock AM, Ringe D, Petsko GA (1992) Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature* 357:423–424.
- Halle B (2004) Biomolecular cryocrystallography: structural changes during flash-cooling. *Proc Natl Acad Sci USA* 101:4793–4798.
- Fraser JS, et al. (2009) Hidden alternative structures of proline isomerase essential for catalysis. *Nature* 462:669–673.
- van den Bedem H, Dhanik A, Latombe JC, Deacon AM (2009) Modeling discrete heterogeneity in X-ray diffraction data by fitting multi-conformers. *Acta Crystallogr D Biol Crystallogr* 65:1107–1117.
- Lang PT, et al. (2010) Automated electron-density sampling reveals widespread conformational polymorphism in proteins. *Protein Sci* 19:1420–1431.
- Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
- Juergens DH, Matthews BW (2004) Cryo-cooling in macromolecular crystallography: Advantages, disadvantages and optimization. *Q Rev Biophys* 37:105–119.
- Sheffler W, Baker D (2010) RosettaHoles2: A volumetric packing measure for protein structure refinement and validation. *Protein Sci* 19:1991–1995.
- Gardberg AS, et al. (2010) Unambiguous determination of H-atom positions: Comparing results from neutron and high-resolution X-ray crystallography. *Acta Crystallogr D Biol Crystallogr* 66:558–567.
- Eisenmesser EZ, et al. (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature* 438:117–121.
- Price AJ, et al. (2009) Active site remodeling switches HIV specificity of antiretroviral TRIMCyp. *Nat Struct Mol Biol* 16:1036–1042.
- Scheidig AJ, Burmester C, Goody RS (1999) The pre-hydrolysis state of p21(ras) in complex with GTP: New insights into the role of water molecules in the GTP hydrolysis reaction of ras-like proteins. *Structure* 7:1311–1324.
- O'Connor C, Kovrigin EL (2008) Global conformational dynamics in ras. *Biochemistry* 47:10244–10246.
- Grant BJ, Gorfe AA, McCammon JA (2009) Ras conformational switching: Simulating nucleotide-dependent conformational transitions with accelerated molecular dynamics. *PLoS Comput Biol* 5:e1000325.
- Wittinghofer A (1998) Signal transduction via Ras. *Biol Chem* 379:933–937.
- Kosloff M, Selinger Z (2003) GTPase catalysis by Ras and other G-proteins: Insights from Substrate Directed Superimposition. *J Mol Biol* 331:1157–1170.
- Scheffzek K, et al. (1997) The Ras-RasGAP complex: Structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* 277:333–338.
- Frauenfelder H, Sligar SG, Wolynes PG (1991) The energy landscapes and motions of proteins. *Science* 254:1598–1603.
- Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, Vendruscolo M (2005) Simultaneous determination of protein structure and dynamics. *Nature* 433:128–132.
- Best RB, Lindorff-Larsen K, DePristo MA, Vendruscolo M (2006) Relation between native ensembles and experimental structures of proteins. *Proc Natl Acad Sci USA* 103:10901–10906.
- Buhrman G, Holzapfel G, Fetis S, Mattos C (2010) Allosteric modulation of Ras positions Q61 for a direct role in catalysis. *Proc Natl Acad Sci USA* 107:4931–4936.
- Heo WD, Meyer T (2003) Switch-of-function mutants based on morphology classification of Ras superfamily small GTPases. *Cell* 113:315–328.
- Matsumoto K, et al. (2011) Critical roles of interactions among switch I-preceding residues and between switch II and its neighboring alpha-helix in conformational dynamics of the GTP-bound Ras family small GTPases. *J Biol Chem* 286:15403–15412.
- Forbes SA, et al. (2011) COSMIC: Mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* 39:D945–950.
- Shaw DE, et al. (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* 330:341–346.
- Lockless SW, Ranganathan R (1999) Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* 286:295–299.
- Mittermaier A, Kay LE (2006) New tools provide new insights in NMR studies of protein dynamics. *Science* 312:224–228.
- Kantardjiev KA, Rupp B (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* 12:1865–1871.
- Voss NR, Gerstein M, Steitz TA, Moore PB (2006) The geometry of the ribosomal polypeptide exit tunnel. *J Mol Biol* 360:893–906.
- Hamelryck T (2005) An amino acid has two sides: a new 2D measure provides a different view of solvent exposure. *Proteins* 59:38–48.
- Lovell SC, Word JM, Richardson JS, Richardson DC (2000) The penultimate rotamer library. *Proteins* 40:389–408.

Supporting Information

Fraser et al. 10.1073/pnas.1111325108

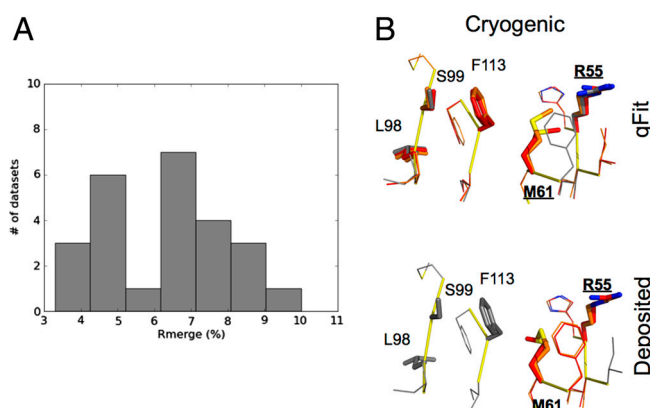


Fig. S1. R_{merge} values and qFit enable comparisons of cryogenic and room-temperature datasets. (A) A histogram of published R_{merge} values for room-temperature datasets indicate that no unusual radiation damage has occurred. All values are below 10% across the entire resolution range. We do not observe widespread effects of radiation damage, such as S oxidation, scission of disulfide bonds, or opening of aromatic rings. Below, we have compiled a list of specific references to the procedures used by the individual investigators to limit radiation damage: *Cyclophilin*: "Room-temperature X-ray data were collected at 15 °C with 96% humidity using a temperature- and humidity-controlled goniometer head (the Rigaku free-mounting device) at ALS Beamline 12.3.1. To limit the effects of radiation damage while maintaining high signal-to-noise, we collected short exposures (0.1 s) for 180 degrees with 1 degree oscillations. A large crystal (1 mm × 0.5 mm × 0.3 mm) was translated halfway through this rapid collection protocol. Following this initial low-exposure pass, an additional complete dataset of 90 degrees with longer exposures (1 s) was collected while translating the crystal every 10 frames. During data processing, we ensured that radiation damage had not degraded the data quality by subdividing the data into subsets of 45 frames and confirming that unit cell parameters, scale factors, and chi-square statistics were consistent throughout the dataset. Owing to the short total exposure time (108 s), the limited exposure of each spot on the crystal, the size of the crystal, and the finite rate of crystal damage, the data were not significantly influenced by radiation damage." (1). *Crystallin*: "A two degree oscillation range was used and the crystal was translated through the beam roughly every 20 degrees in order to minimize radiation damage." (2). *Pilin*: "To obtain the final dataset, scaling and merging were carried out in a non-standard manner. Firstly, batch 1 was scaled and merged separately. Next all four batches were processed together and scaled to batch 1. Finally, the program SFTOOLS was used to create the missing reflections from the scaled set of all data. This minimized the effect of radiation damage without compromising the data completeness" (3). *RNAse T1*: "Their intensities were corrected for crystal decay, which did not exceed 9%" (4). *Rubredoxin*: "Each frame consisted of a 1 degree oscillation with .5 s exposure time. To counter the heating effect of the X-ray beam, the beam was attenuated to 25% of full strength and a 295 K gas stream was flowed over the crystal." (5). *Xylanase*: "To cover the whole range of intensities, three sets of images were recorded, differing in exposure times and resolution limits, for the cryodata and two sets of images were recorded for the room-temperature data." (6). (B) Model building and refinement of CypA using qFit shows that the cryogenic X-ray data reveal a single conformation for many residues (Leu98, Ser99 and Phe113) that agree with the deposited model.

- 1 Fraser JS, et al. (2009) Hidden alternative structures of proline isomerase essential for catalysis. *Nature* 462:669–673.
- 2 Najmudin, et al. (1993) Structure of the bovine eye lens protein gammaB(gammall)-crystallin at 1.47 Å. *Acta Crystallogr D Biol Crystallogr* 49:223–233.
- 3 Dunlop, et al. (2005) Pros and cons of cryocrystallography: Should we also collect a room-temperature data set? *Acta Crystallogr D Biol Crystallogr* 61:80–87.
- 4 Martinez-Oyanedel, et al. (1991) Ribonuclease T1 with free recognition and catalytic site: Crystal structure analysis at 1.5 Å resolution. *J Mol Biol* 222:335–352.
- 5 Gardberg AS, et al. (2010) Unambiguous determination of H-atom positions: Comparing results from neutron and high-resolution X-ray crystallography. *Acta Crystallogr D Biol Crystallogr* 66:558–567.
- 6 Natesh, et al. (2003) Thermostable xylanase from *Thermoascus aurantiacus* at ultrahigh resolution (0.89 Å) at 100 K and atomic resolution (1.11 Å) at 293 K refined anisotropically to small-molecule accuracy. *Acta Crystallogr D Biol Crystallogr* 59:105–117.

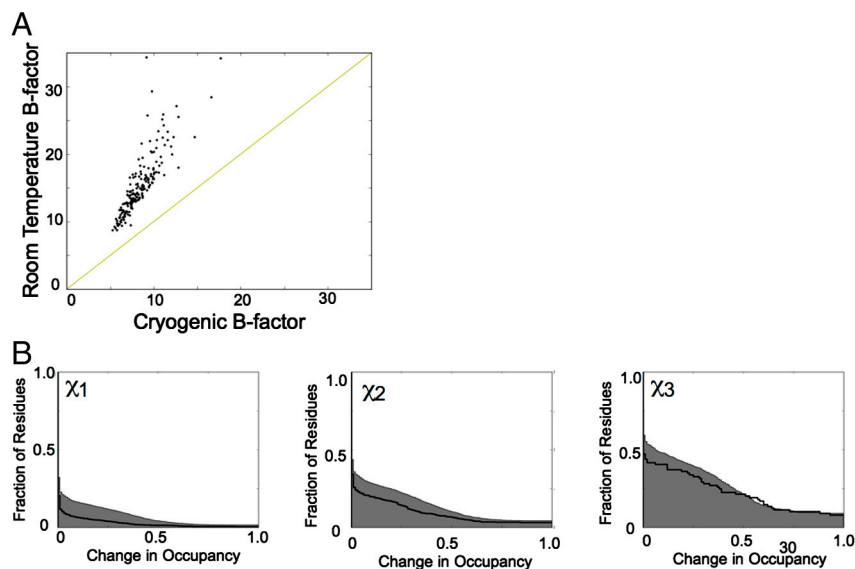


Fig. S2. Changes in B-factors and rotamers induced by crystal cryocooling. (A) A scatter plot of residue-averaged B-factors for cryogenic and room-temperature xylanase II. Although crystal cryocooling causes only modest structural effects on the overall folds of all 30 proteins, the average atomic B-factors are reduced and the B-factor distributions are narrowed compared to models refined against isomorphous room-temperature data. Residues that deviate from predicted linear reductions in B-factors may reflect local variations in effective temperature, differential coupling to the solvent or the lattice, or changes in the relative populations of alternative conformations. Rather than directly modeling the populations of interconverting alternative conformations, the B-factors approximate local harmonic displacements. All residues fall above the yellow equality line, which indicates that B-factors capture increased thermal motion at room temperature. (B) Cumulative histograms of the maximum absolute change in rotamer occupancy from qFit models for all residues (shaded gray area) and buried residues (black line) reveal that cryocooling affects the conformational substates at multiple χ -angles.

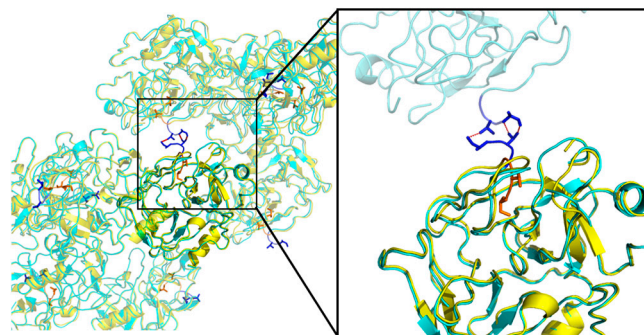


Fig. S3. A functionally important loop in trypsinogen forms new lattice contacts upon crystal cryocooling. In trypsinogen, which has the largest rmsd of any structure pair in our dataset (1.1 Å), cryocooling induces a conformational change in the N-terminal terminal loop. The room-temperature (yellow) and cryogenic (cyan) models of trypsinogen are shown in the crystal lattice environment. The inset focuses on the N-terminal loop, which adopts a buried conformation at room temperature (yellow/orange). In contrast, at cryogenic temperatures, the loop adopts an exposed conformation and forms new symmetric side-chain lattice contacts (red dashes) in a solvent channel. This loop is flexible in solution and is important for the proteolytic maturation into trypsin. Trypsinogen provides a dramatic example of how cooling causes small but significant structural changes and the formation of additional lattice contacts. The differences between the room-temperature and cryogenic structures of trypsinogen highlight functionally important flexible residues that are mobile in solution.

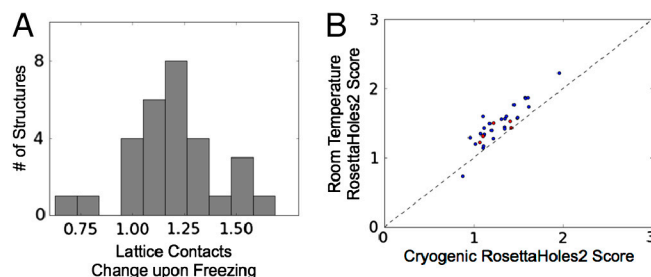


Fig. S4. Cryocooling changes lattice contacts and protein packing. (A) A histogram of the change in lattice contacts indicates that cryocooling increases lattice contacts. Lattices with increased contacts (cryogenic/room-temperature ratio >1.0) may help to stabilize the arrangement of proteins in the crystal, leading to increases in resolution at cryogenic temperatures. (B) RosettaHoles2 packing score for room-temperature vs. cryogenic models. Lower scores indicate better packing. All pairs of refined models, except partially deuterated rubridoxin, show quantitatively superior packing at cryogenic temperatures. Proteins where the room-temperature dataset is higher resolution are shown in red.

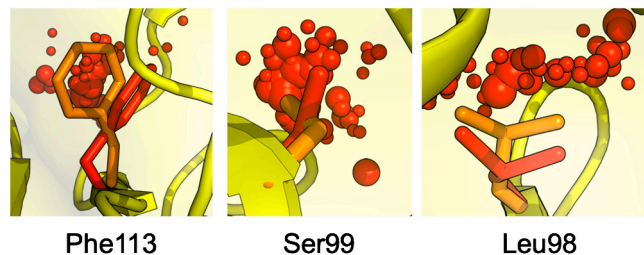


Fig. 55. The association of packing defects and alternative conformations. Packing defects (red spheres) are partially filled by nearby alternative conformations (orange) of Phe113, Ser99, and Leu98 in CypA. Only the major conformations (red) are detectable by Ringer or qFit at cryogenic temperatures, where these packing defects are absent.

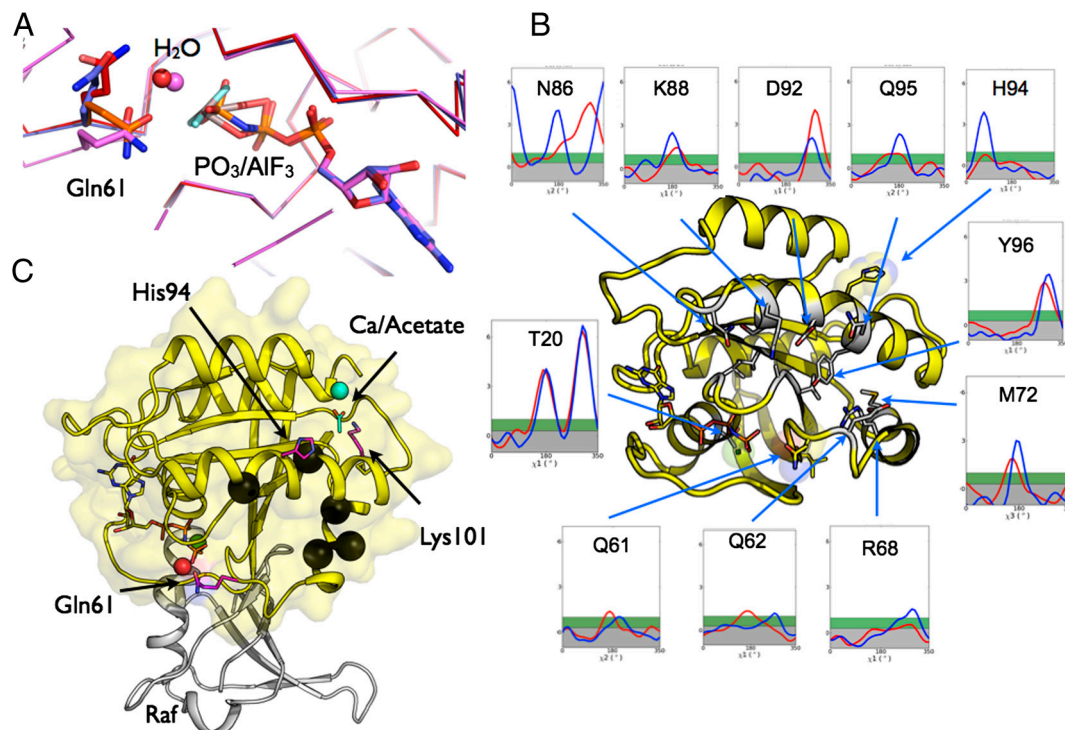


Fig. 56. Temperature dependence of the H-Ras:GTP allosteric mechanism. (A) Substrate-based superposition of H-Ras cryogenic (blue), room-temperature (red and orange), and transition state (violet) models (as in Fig. 5B). (B) The room-temperature qFit model and Ringer plots link helix 3 residues to the catalytically competent active site by temperature-dependent changes in conformational distributions. (C) Perturbations to the dynamic network, which extends from the cavity at the C-terminal end of helix 3 to the active site, identified by relaxation dispersion NMR in H-Ras and room-temperature electron-density sampling, can cause long-range changes in the conformational ensemble and activity of H-Ras. However, each perturbation selects from the ensemble different conformations with distinct functional properties. Recent X-ray studies have revealed that binding of Ca^{2+} and acetate (shown in cyan, upper right) in the allosteric cavity modulates the conformational ensemble (from PDB ID code 3LBH). Although Ca^{2+} is not known to directly regulate H-Ras, Ca^{2+} binding in the crystal prevents collapse of the cavity, selects alternative side-chain conformations along helix 3, and stabilizes the catalytically competent conformation of Gln61 (shown in pink lower left). The key residues implicated in this conformational change match flexible residues identified in the absence of Ca^{2+} both by NMR relaxation and by our analysis of the room-temperature ensemble. Mutation of Lys101 (shown in pink) near the allosteric cavity alters the morphological outcome of interactions between Raf (shown in gray from PDB ID code 3KUD) and the Ras active site, including Gln61 (shown in pink). Mutations identified in cancers are shown as black spheres adjacent to helix 3 and His94 (shown in pink).

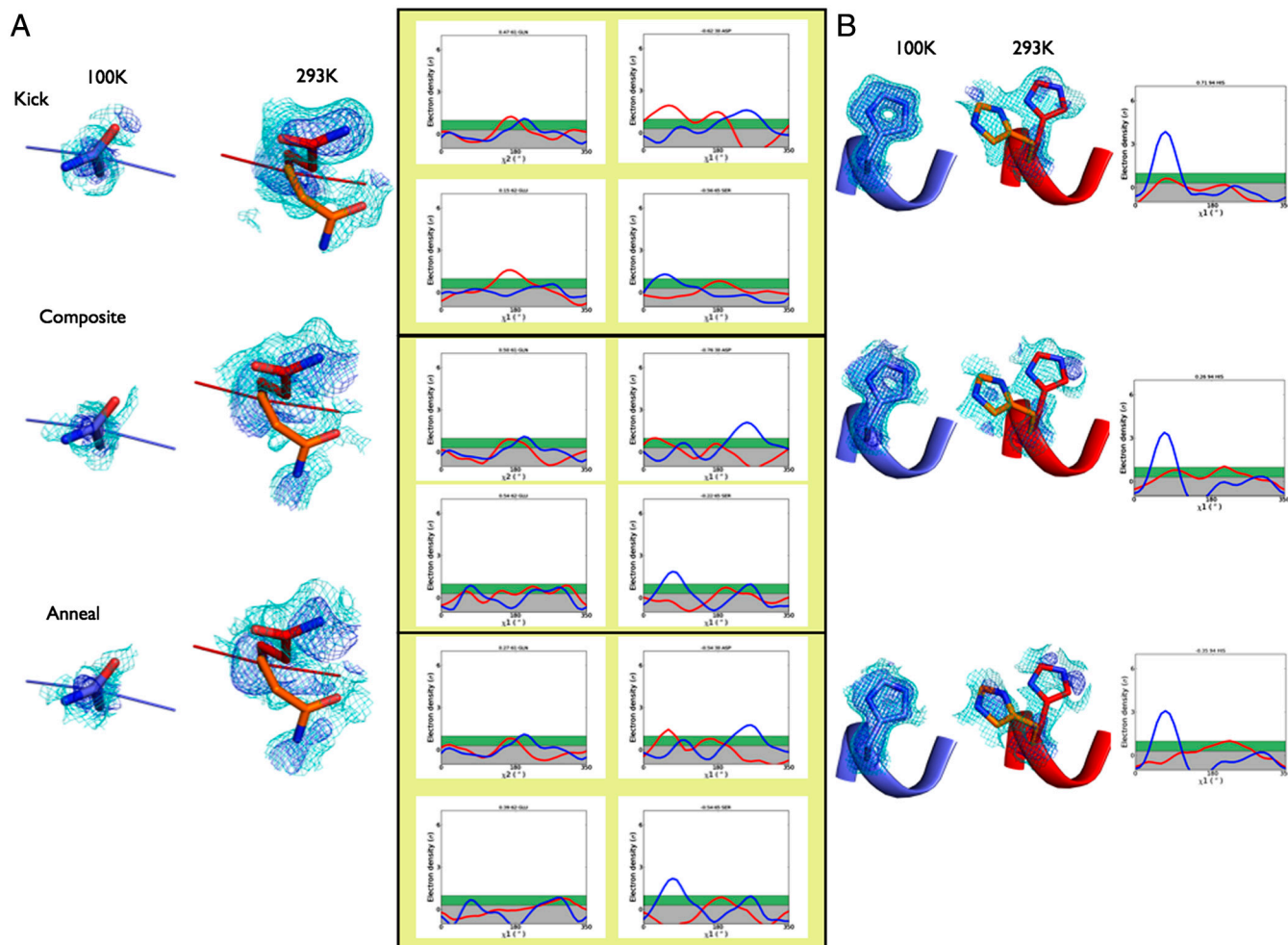


Fig. S7. Minor populations are detected in different types of electron-density maps. (A) A comparison of Phenix-kicked, composite-omit, and simulated-annealing omit maps for the active site residues of H-Ras reveals electron density consistent with a second conformation of Gln61. Electron-density distributions reveal differences between the room-temperature and cryogenic maps in these map types, consistent with the interpretations made using maximum-likelihood weighted maps. The reduction in the -60° peak of Gln61 at χ_2 is due to a coupled backbone/side-chain shift that places the peak corresponding to the alternative C δ position outside the sampling radius of Ringer. (B) Different electron-density maps for His94 of H-Ras reveal features consistent with a second conformation at room temperature. Ringer plots reveal differences between the room-temperature and cryogenic maps in all three map types, consistent with the interpretations made using maximum-likelihood weighted maps. The differences in the H-Ras active site could be due to factors other than temperature (for example, a “bad dataset”). However, there are several considerations that argue against this possibility. The cryogenic model refined well (Table S1), and the map is of very high quality in other regions, for example His94. Analysis of the reflections by phenix.xtriage indicates no indicators of dataset pathologies such as: twinning, ice-rings, or anisotropy. This suggests that the difficulty in interpreting the Gln61 area is not a property of the crystal or the data quality but rather stems from the cryogenic procedure. The room-temperature map, in contrast, can be interpreted as exchanging between the “active” and “inactive” conformation. While the cryogenic data are best interpreted as a single conformation with high B-factors, the room-temperature electron-density maps reveal multiple conformations with different potential activities. This difference suggests that cryogenic artifacts may complicate the interpretation of biological mechanisms in weak electron density.

