BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

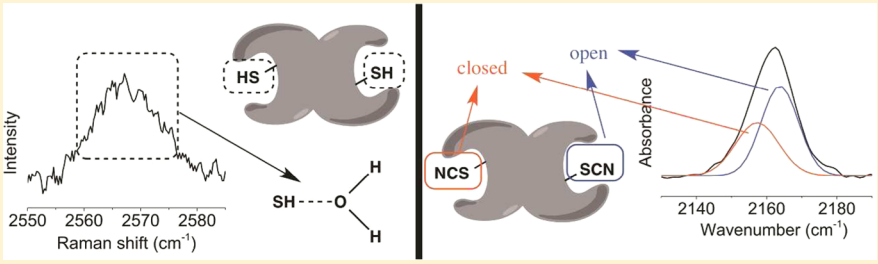
pubs.acs.org/biochemistry

Dynamic Asymmetry and the Role of the Conserved Active-Site Thiol in Rabbit Muscle Creatine Kinase

Casey H. Londergan,* Rachel Baskin, Connor G. Bischak, Kevin W. Hoffman, David M. Snead, and Christopher Reynoso

Department of Chemistry, Haverford College, 370 Lancaster Avenue, Haverford, Pennsylvania 19041, United States

Supporting Information



ABSTRACT: Symmetric and asymmetric crystal structures of the apo and transition state analogue forms, respectively, of the dimeric rabbit muscle creatine kinase have invoked an "induced fit" explanation for asymmetry between the two subunits and their active sites. However, previously reported thiol reactivity studies at the dual active-site cysteine 283 residues suggest a more latent asymmetry between the two subunits. The role of that highly conserved active-site cysteine has also not been clearly determined. In this work, the S−H vibrations of Cys283 were observed in the unmodified MM isoform enzyme via Raman scattering, and then one and both Cys283 residues in the same dimeric enzyme were modified to covalently attach a cyano group that reports on the active-site environment via its infrared CN stretching absorption band while maintaining the catalytic activity of the enzyme. Unmodified and Cys283-modified enzymes were investigated in the apo and transition state analogue forms of the enzyme. The narrow and invariant S−H vibrational bands report a homogeneous environment for the unmodified active-site cysteines, indicating that their thiols are hydrogen bonded to the same H-bond acceptor in the presence and absence of the substrate. The S−H peak persists at all physiologically relevant pH's, indicating that Cys283 is protonated at all pH's relevant to enzymatic activity. Molecular dynamics simulations identify the S−H hydrogen bond acceptor as a single, long-resident water molecule and suggest that the role of the conserved yet catalytically unnecessary thiol may be to dynamically rigidify that part of the active site through specific H-bonding to water. The asymmetric and broad CN stretching bands from the CN-modified Cys283 suggest an asymmetric structure in the apo form of the enzyme in which there is a dynamic exchange between spectral subpopulations associated with water-exposed and water-excluded probe environments. Molecular dynamics simulations indicate a homogeneous orientation of the SCN probe group in the active site and thus rule out a local conformational explanation at the residue level for the multipopulation CN stretching bands. The homogeneous simulated SCN orientation suggests strongly that a more global asymmetry between the two subunits is the cause of the CN probe's broad and asymmetric infrared line shape. Together, these spectral observations localized at the active-site cysteines indicate an intrinsic, dynamic asymmetry between the two subunits that exists already in the apo form of the dimeric creatine kinase enzyme, rather than being induced by the substrate. Biochemical and methodological consequences of these conclusions are considered.

Many proteins homo- or hetero-oligomerize to form supramolecular complexes with new cooperative function. Understanding the design principles behind the formation of such assemblies could help to eliminate, augment, or mimic the function of oligomeric or supramolecular protein complexes. Creatine kinase (CK) has been noted as a classic[1,2] biochemical example of such a protein in that it functions as a dimer,[3−6] with crystallographic studies suggesting negative cooperativity between the two subunits.[7−9] The asymmetry observed in crystallized transition state analogue complex (TSAC) structures, with substrate present in only one of the two active sites of the dimer, has been ascribed to an "induced fit" mechanism. Such structural asymmetry and apparent intersubunit cooperation could be central to CK's catalytic function and substrate turnover and may be generalizable with respect to other proteins that feature asymmetric or half-of-sites cooperativity (originally proposed by Koshland[10]) in their functional oligomeric complexes. This article presents a new view of the dynamic structure of CK using novel analytical

Special Issue: New Frontiers in Kinases

Received: July 1, 2014
Revised: October 1, 2014
Published: October 27, 2014

ACS Publications © 2014 American Chemical Society

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

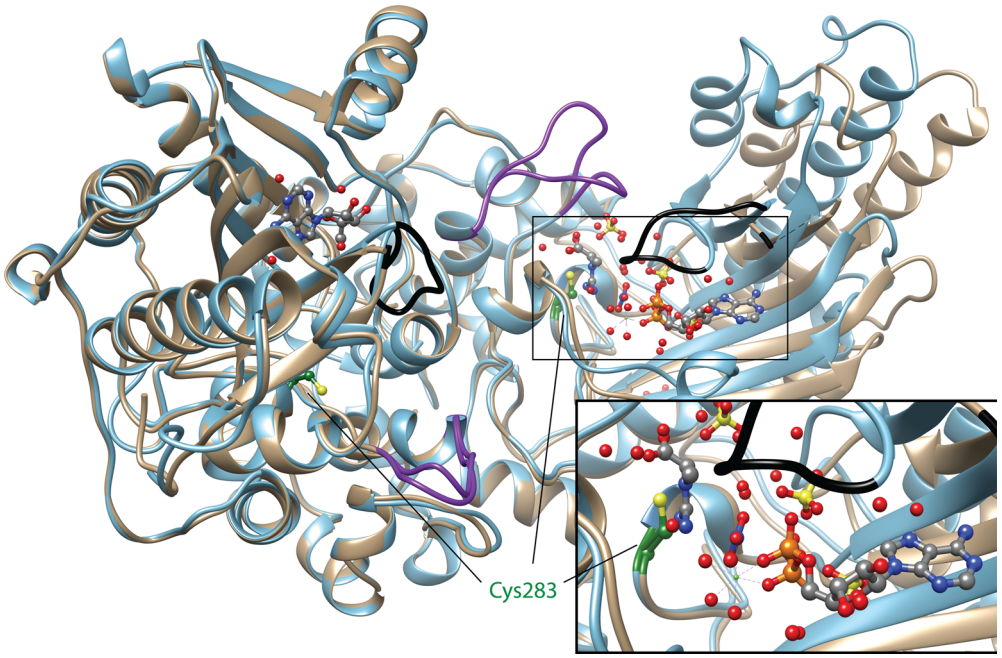


Figure 1. Overlaid crystal structures of the apo-rmCK dimer (from the structure of Protein Data Bank entry 2CRK,[12] colored blue) and the TSAC structure of the R134K mutant (1U6R,[8] colored tan). The inset shows a close-up of the boxed area in the subunit at the right. The "left" subunit of the dimer is nearly identical within crystallographic resolution, while the "right" subunit differs substantially because of the presence of all substrates in the active site of that subunit. Cys283, the subject of this study, is shown as balls and sticks and colored green with the sulfur atom colored yellow. All substrates, cofactors, and crystallographic water molecules are represented with balls and sticks; ADP appears at the right of the right active site as well as in the left active site, while creatine appears close to Cys283 only in the right active site. Loops previously implicated in possible motions that affect the substate accessibility of the active site are colored purple (the 62−70 loop) and black (the 322−330 loop), with a substantial difference observed for the purple loop between the apo and TSAC structures of the right subunit.

techniques focused on functionally important residues in the dual active sites.

CK, present in all vertebrates, belongs to a large family of phosphagen kinases that transfer a phosphoryl group from ATP onto a guanidinium group. Other examples of phosphagen kinases include arginine kinase (AK), glycocyamine kinase, taurocyamine kinase, and lombricine kinase. Any new conclusions reached about the catalytic mechanism of CK might also apply to these other structurally and functionally homologous enzymes. CK catalyzes the reversible transfer of γ-phosphate of MgATP to creatine to produce MgADP, a proton, and phosphocreatine (PCr). PCr serves as an energy reservoir that supplies ATP under stress conditions, and the participation of CK in its production has led to popular assay-based diagnostic tests of stress conditions in humans undergoing cardiac arrest. CK is a central enzyme in cellular energy metabolism and is found in organs with high metabolic activity such as the heart, brain, and muscle. CK's sequence varies throughout vertebrates with structural homology between CK's from different tissues and species. Three 85 kDa homodimeric isoforms are found in mammals: brain type BB, muscle type MM, and the MB heterodimer, which are all functional as dimers, along with two octameric mitochondrial isoforms that are not the subject of this work.[11] The amino acid sequences of these isozymes consist of six conserved regions and seven variable regions. Although the structure and catalytic mechanism of CK have been studied for many decades, a complete understanding of some conserved structural elements and the intersubunit relationship remains elusive.

Kinetic analyses of several CK's have shown that the phosphorylation reaction is reversible and operates via a rapid equilibrium random bimolecular mechanism *in vitro* above pH 8.0 following a well-understood chemical mechanism.[6] The efficiency of the CK enzymatic process is consistent for most forms of CK. Typical k_{cat} values range from 3×10^3 to 9×10^3 min^{-1}, whereas typical K_M values for creatine binding range from 6×10^2 to 3×10^3 μM.[6]

Several crystal structures of varying CK's have been published, including high-resolution structures for rabbit muscle CK (rmCK),[12] chicken brain CK,[13] and the human ubiquitous mitochondrial isozyme.[14] These crystal structures are all strongly homologous, especially at the twin active sites of the dimeric enzyme. The reported crystal structures of apo CK dimers all exhibit exact structural symmetry between the subunits.[12−14] CK crystal structures have helped to elucidate which specific residues are important in linking the two subunits, and all of the residues participating in the active sites have been identified. On the basis of crystal structures and sequence homology, there are two highly conserved sequences that might play a large role in the enzymatic mechanism: a negatively charged "NEED box" and a region that surrounds a free cysteine residue that is reactive to solution-phase thiol-modifying reagents. The free cysteine is highly conserved across all vertebrate species and also appears in invertebrate AK enzymes,[15] so the thiol is presumed to be important to the catalytic function of the enzyme. The reported apo and TSAC structures of rmCK are shown superimposed in Figure 1, with an emphasis on the active site and its reactive thiol group.

The reactive and conserved cysteine (Cys283 in rmCK) in the active site of CK has been the subject of extensive study,[5,16] but the exact catalytic role of this residue remains unclear. According to previous work, chemical modifications of this cysteine's thiol generally either decrease the activity of CK or inactivate the enzyme completely. Additionally, several thiol-

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

modifying reagents differentially modify one of the two active sites of the dimer, thus strongly suggesting that there is an intrinsic dimeric asymmetry that extends to the reactivity of the two thiols.[17−22] Measured enzymatic activity was decreased but not abolished upon mutation of Cys283 to either glycine, serine, alanine, asparagine, or aspartate;[23] therefore, the Cys283 residue in rmCK (or the homologous Cys in other CK's) is apparently optimal but not essential for catalysis or substrate binding. It was suggested that the cysteine residue might play a vital role in synergistic substrate binding, in which the binding of the first substrate (i.e., ATP) allows additional substrates (i.e., creatine) to bind with greater affinity to the CK active site before the reaction proceeds.[23] In this context, Cys283 would hypothetically help to facilitate a local conformational change upon substrate binding by shaping the active site of CK to promote catalysis.

More recent studies examined the pK_a of the analogous Cys282 of human muscle CK (hmCK). An initial spectrophotometric study suggested that the pK_a of Cys282 is 5.6, which is much lower than the standard pK_a of 9.1 for free cysteine residues,[24] and further computational studies of hmCK predicted that the active-site cysteine should be a deprotonated thiolate ion at physiological pH.[25] However, residual enzymatic activity remains upon modification of the thiol of the active cysteine to an uncharged form (see below), suggesting that the thiol might actually be protonated at catalytically relevant pH's.[17,18]

Despite the observed differential reactivity to several Cys-modifying reagents, and overwhelming evidence leading to the consensus that most CK's function as dimers, the importance of dimerization to the catalytic mechanism of CK has remained unclear.[26] Kinetic studies of monomeric rmCK mutants whose intersubunit interface was intentionally disrupted showed residual enzymatic activity, though some remaining dimerization may have affected those results.[27] Outside of that single study, it has generally been concluded that rmCK's enzymatic activity is directly connected to its dimeric quaternary structure: when dissociated from the dimer conformation, individual monomers usually become unstable and have a high propensity for unfolding.

The more recent crystallographic work[7,8] suggested that the two dimeric enzyme active sites exhibit negative cooperativity, where substrate binding to one active site renders the other active site inactive. Methylation of both active-site cysteines led to ~20% (and notably not zero) enzymatic activity. Double-reciprocal plots for the CH_3S-blocked enzyme with MgATP were biphasic, suggesting the continuing possibility of negative cooperativity in the metal−nucleotide binding step in the thiomethyl-slowed enzyme.[19]

The TSAC crystal structure of rmCK[8] in the presence of ADP, Mg^{2+}, creatine, and NO_3^{2-} is asymmetric, in contrast to the apo crystal structure of rmCK, which consists of two subunits with identical structures. This structural difference in only one of the subunits is exhibited graphically in Figure 1. In the TSAC structure, one subunit is "closed" with a fully formed TSAC while the other subunit contains only MgADP in the active site. The conformational changes in two loop regions in only one monomer appear to correlate with negative cooperativity upon binding, where creatine can bind to only one active site per dimer at a time. Differences in loop structures associated with "closing" the active site between the apo and TSAC structures (see these loops in Figure 1) imply a possible mechanism of communication between the two subunits upon substrate binding.[7,8] TSAC structures of CK's from human and snail show asymmetry similar to that observed in rmCK,[7,9] so 100% negative cooperativity between the twin active sites (or "half-of-sites reactivity") appears to be a general feature of this family of enzymes.

The TSAC crystal structures that have been used suggest that intersubunit cooperativity in CK dimers is induced by substrate binding, and this idea leads to a picture of CK catalysis in which the structures of the two active sites are controlled by a mutual induced fit mechanism that confers the negative cooperativity. The induced fit argument is made solely on the basis of the differences between the apo and TSAC crystal structures. Although the crystal structure of the apo form of rmCK is symmetric, prior cysteine reactivity studies by Degani et al.[17,18] showed that the two active-site cysteines of the apo rmCK dimer react differently in solution with some thiol-modifying reagents (summarized in Figure 2, with NTCB as a primary

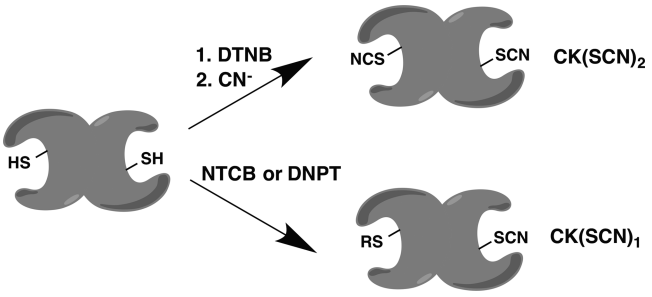


Figure 2. Cysteine cyanylation reactions performed on rmCK by Degani and co-workers.[17,18] This graphic follows their previously stated conclusions.[5]

example of a reagent that modifies only one of the two thiols), suggesting that the subunits of dimeric rmCK might be arranged asymmetrically in solution in the all-apo form before substrate binding. There is no structure-based explanation in the literature for why NTCB might modify only one subunit, and this work hopes to provide some clarity on this point. On the basis of the catalytic activity following different thiol modifications, Degani postulated that the two subunits of rmCK function cooperatively with communication between the two active sites, but following a mechanism not dependent on substrate binding. There is a fundamental disagreement between the substrate-induced asymmetry explanation based on the TSAC crystal structure and the asymmetry between the two active sites in thiol reactivity to certain reagents. A goal of this work is to clarify the structural and/or dynamic basis for both of these observed asymmetries in the rmCK dimer.

To unify the more recent structure-based crystallography results with thiol reactivity and activity assays, a solution-phase structure-based technique must be implemented. In such a large dimeric protein, nuclear magnetic resonance (NMR) spectroscopy is not expected to be feasible to determine the complete structure. However, vibrational spectroscopy is an ideal technique for examining a small part of a larger structure in proteins of arbitrary size and in many different media,[28] especially in cases where the small region of interest can be isolated from the rest of the protein in the vibrational spectrum by site-specific labeling.[29,30]

The S−H vibration of cysteine appears between 2500 and 2600 cm^{-1} and falls in the "clear" spectral window between 1900 and 2700 cm^{-1} that is devoid of vibrations from other

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

biomolecular functional groups. The S−H band is a weak and inconsistent IR absorber because of the covalent nature of the S−H bond and the polarizability of sulfur;[31,32] however, these features make the S−H stretching mode a strong and easily observable Raman scattering chromophore.[33] The S−H frequency depends strongly on the participation of the S−H group in hydrogen bonds.[34] In bulk-type liquid water, the S−H frequency is ∼2575 cm^{-1}; if the thiol acts as a H-bond donor, the S−H band frequency decreases, and if the sulfur lone pairs accept a H-bond, the S−H band frequency increases. Model compound studies in organic solvents indicate that the S−H band can report the thiol's acting both as a H-bond donor and as a H-bond acceptor with amides[34] (see also additional results in the Supporting Information). The S−H central frequency is not very sensitive to local dipoles, according to our own model compound work (see the Supporting Information). For free cysteine residues, the presence and relative intensity of the S−H band in the Raman spectrum can also be used to determine the protonation state of the S−H moiety without any spectroscopic interference from other parts of the protein.

Spectroscopic isolation of a specific site can also be achieved through the incorporation of an artificial amino acid side chain at the site of interest. β-Thiocyanato-alanine, or cyanylated cysteine (C*), has a clear CN stretching vibration near 2160 cm^{-1} that appears in the same vacant spectral window between 1900 and 2700 cm^{-1}. The thiocyanate CN stretching frequency is sensitive to both the local dipolar environment and the presence of H-bond donors.[35,36] Strong local dipolar fields decrease the CN frequency, while H-bond donors (most notably H_2O) increase the CN frequency. The CN line shape is also able to report on the picosecond dynamics and the local dipolar and H-bonding structural distribution resolved on the picosecond time scale.[36] Given this complicated dependence of the CN frequency on the SCN probe group's environment, quantitative interpretation of the CN band's frequency and line shape in enzymatic active sites and at a protein−protein binding interface has been challenging.[37−41] However, it is quite clear from studies in model peptides[42,43] and disordered proteins[44] that if the SCN probe group is completely solution-exposed in near-bulk-like H_2O, its frequency is 2163 cm^{-1}, while as the nitrile becomes completely non-water-exposed, its frequency drops to 2154−2159 cm^{-1}. This particular sensitivity to the local presence of water has led to C*'s use in identifying, for example, the membrane-bound side chains of amphiphilic peptides in contact with lipid membranes.[42,45] C* is an ideal active-site probe for rmCK because it has already been demonstrated that C* can be incorporated at both Cys283 sites (via sequential reaction with DTNB and cyanide) or only one site (via reaction with NTCB, as in Figure 2).[18]

The first goal of this work is to determine the chemical environment and role of the conserved active-site cysteine residue using vibrational spectroscopy. The second, more general goal is to examine how the environments of the active sites of rmCK vary in each subunit, with the goal of identifying any clear asymmetry between the two subunits in the absence and presence of CK's substrate and cofactors. First, we observed S−H stretching vibrations from unmodified Cys283 side chains to determine the dynamic environment of the Cys283 residue. The unmodified rmCK proteins were subjected to both varying pH and conditions favorable to the formation of the TSAC. Next, to determine the dynamic environment inside each active site from the point of view of the free Cys283 residue, we converted the rmCK dimer to either $CK(SCN)_2$ or $CK(SCN)_1$ using the cyanylation procedures described by Degani et al., and we used infrared spectroscopy of the CN stretching band(s) to report on the microenvironments of probe groups in the two active sites. The active $CK(SCN)_2$ protein was also placed under TSAC-forming conditions. Finally, we used molecular dynamics simulations of the unmodified and Cys-modified proteins to help interpret the spectroscopic data from both unlabeled and CN-labeled proteins.

MATERIALS AND METHODS

Materials. Sodium cyanide (NaCN), dithiothreitol (DTT), creatine, ADP, ATP, *p*-enolpyruvate, nicotinamide adenine dinucleotide (NaDH), HEPES, potassium chloride, EDTA, glacial acetic acid, tris base, and pyruvate kinase were purchased from Aldrich. Monobasic phosphate and magnesium acetate were purchased from J. T. Baker. Dibasic phosphate and sodium nitrate were purchased from Fischer. Lactate dehydrogenase was purchased from Calbiochem. All chemicals mentioned above were used as received. All buffers were made with doubly deionized H_2O and degassed using a Millipore 0.20 μm nylon membrane prior to use. rmCK (MM isoform) was obtained from USB Chemical as a lyophilized powder and purified as described in the Supporting Information. In all experiments, the rmCK concentration was determined using an extinction coefficient of 71000 cm^{-1} M^{-1} at 280 nm as originally reported.[1]

Raman Spectroscopy. All Raman spectra were recorded at 22 °C using a home-built Raman spectrometer, reported schematically here and in more detail elsewhere.[46] Samples in glass capillaries housed in a brass temperature-control block were excited using the 514.5 nm line of a Coherent Innova 307C Ar^+ laser at an incident power of approximately 50 mW. Scattered light at 90° from the excitation axis was collected via a Nikon Nikkor f/2.2 lens, collimated, focused into and separated using a PI-Acton SpectraPro 2500i 0.5 m single-pass monochromator, and imaged using a Spec-10/300 liquid N_2-cooled CCD array, also from PI-Acton. Spectra of unmodified CK were collected for 60 min using a 600 grooves/mm grating blazed at 500 nm (for the pH titration curve) or for a total of 24 h using a 2400 grooves/mm grating (noted below as "high-resolution" spectra for line shape analysis of the TSAC vs apo forms) and were the sum of 10−30 consecutive shorter exposures.

Cyanylation of rmCK. rmCK was cyanylated at Cys283, the only solvent-exposed, free cysteine residue, via two methods described by Degani et al.[18,47,48] The reactions were typically performed on a scale of 15−20 mg of purified protein (240 nmol of dimer) in 500 μL of solution in a semimicro quartz UV cell with a path length of 5 mm (Starna Cells). The buffer for all reaction mixtures and stock solutions was 0.5× TAE (pH 7.0). See the Supporting Information for a time-based quantification of the modification yield from each procedure.

$CK(SCN)_2$. To a stock solution of rmCK was added DTNB in an 8:1 molar excess from a 25 mM DTNB stock solution in 0.5× TAE (pH 7.0). After 30 min, the reaction yield was measured via a Jasco V-570 UV/vis/NIR spectrophotometer based on the release of TNB, with an extinction coefficient for TNB of 14150 cm^{-1} M^{-1} at 412 nm.[49,50] Sodium cyanide (NaCN) was then added at a 55:1 molar excess, leading to cyanylation of all TNB−protein adducts and all remaining DTNB. After 30 min, excess reagents were removed via size exclusion chromatography on a PD-10 column packed with

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

Sephadex G-25 (GE Healthcare) with 0.5× TAE (pH 7.0) used as an elution buffer. Fractions containing creatine kinase were collected and concentrated to ∼20 μL using Amicon Ultra 10k centrifugal filter devices.

CK(SCN)$_1$. 2-Nitro-5-thiocyanobenzoic acid (NTCB) was synthesized and isolated as a solid using a previously described procedure (details are provided in the Supporting Information). NTCB was then redissolved to a concentration of 10 mM in 0.5× TAE buffer and added in 3−9-fold excess to rmCK. The reaction yield was monitored via the TNB side product at 412 nm, as described above. Nonprotein reaction products were removed as described above, and the resulting protein was concentrated to a volume of ∼20 μL.

Infrared Spectroscopy. All infrared spectra were collected at room temperature in a 22 μm path length BioCell (BioTools, Inc., Jupiter, FL) using a Vertex 70 FTIR spectrometer (Bruker Optics). IR spectra of protein samples were collected using 1024 averaged scans at 2 cm^{-1} resolution. A background spectrum of the appropriate buffer was subtracted from each sample spectrum, and the baseline for the CN stretching region was additionally corrected by fitting the baseline (excluding the region between 2140 and 2190 cm^{-1}) to a polynomial and subtracting the fit. Line shape analysis and fitting were conducted using Origin versions 7 and 8.

TSAC Formation. Twenty-five to thirty milligrams of creatine kinase (either purified and unmodified or doubly cyanylated using the first procedure described above) was obtained. The buffer was then exchanged twice for a mixture of 6.25 mM magnesium acetate, 25 mM creatine, 12.5 mM sodium nitrate, and 6.25 mM ADP in 25 mM HEPES buffer (pH 7.0) using a PD-10 column packed with Sephadex-25. The protein in the resulting solution was concentrated using Amicon Ultra 10k centrifugal filter devices. The final protein concentration of the unmodified Raman samples was ∼1 mM, and the final protein concentration of the IR sample of $CK(SCN)_2$ was 1−2 mM; each of these protein concentrations is lower than the solution concentration from which the rmCK TSAC was crystallized.[8]

Molecular Dynamics Simulations. GROMACS[51] 4.6.1 was used to perform all molecular dynamics simulations. Starting Protein Data Bank (PDB) structure files came from the RCSB Protein Data Bank (2CRK[12] for all-apo rmCK and 1U6R[8] for the TSAC of rmCK). Chimera UCSF[52] was used to add the SCN probe to the desired residues and to make a dimer of the CK monomer in the symmetric apo crystal structure. PyMOL (Schrodinger, Inc.) was used to achieve the desired starting dihedral angles for different starting orientations of the Cys283 residue. Profix, a part of Jackal (Honig laboratory, Columbia University, New York, NY), was used to fix missing residues or atoms in the PDB file. Visual Molecular Dynamics (VMD, Schulten laboratory, University of Illinois at Urbana-Champaign, Urbana, IL) was used to visualize the results of trajectories on a frame-by-frame basis from trajectory files and to perform angle, bond distance, and hydrogen bond analysis. Structural graphics were made using VMD or Chimera. All GROMACS preparation steps were performed on local linux workstations, and MD trajectories were run on Fock, an in-house linux cluster at Haverford College, using single multicore nodes.

The Amber99SB[53] force field was used with explicit TIP3P[54] water. Additional force field parameters to describe the SCN functional group[55] were kindly provided by J. Layfield and S. Hammes-Schiffer (University of Illinois at Urbana-Champaign, Urbana, IL), based on prior quantum chemistry calculations on the SCN group by Cho's group.[56,57] The SCN group's initial geometry was set using results of a B3LYP[58]/6-311G(d,p) minimization of ethyl thiocyanate in Gaussian03.[59]

Simulations were performed in a cubic box with periodic boundary conditions and a ≥1.0 nm separation between the protein and the box's sides. All initial CK structures were energy-minimized using the steepest descent algorithm until the maximal force was <1000 kJ/mol/nm. Then an *NVT* ensemble equilibration was performed for 100 ps with 2 fs time steps until a plateau at the desired temperature of 300 K was reached. All water molecules were placed using the genbox command in Gromacs and were not necessarily at any crystallographically specified positions; the equilibration routine was robust enough that active-site water molecules should not be stuck in structures with strong local minima. For fixed SCN orientation runs, the entire Cys283(SCN) amino acid was held fixed throughout the equilibration. Production trajectories of 5 ns were generated with 1 fs time steps.

RESULTS AND DISCUSSION

Raman S−H Stretching Bands of Unmodified rmCK. We used differences in the S−H stretching region between the Raman spectra of the doubly cyanylated (at Cys283, from which the Cys283 S−H stretching bands should be absent) and unmodified forms of the protein to assign the Cys283 signals in a manner separate from other free cysteines in the sequence. There are three other free cysteines buried in the folded protein whose signals should be invariant to cyanylation; these Cys residues display much broader signals at frequencies lower than that of Cys283, whose signal greatly diminishes between the cyanylated and unmodified forms of the protein. See the Supporting Information for more details about the process used to assign the sharp S−H signal discussed below to Cys283.

The detailed S−H stretching Raman scattering bands for Cys283 in the unmodified, apo rmCK and the TSAC of the unmodified enzyme at pH 7.0 are shown in Figure 3. There are no significant differences between the two spectra or their fitted line shapes, which are completely symmetric and quite narrow as compared to S−H bands observed elsewhere,[33,34,60−63] in our model compound results (see the Supporting Information), and also as compared to the other S−H bands from this protein

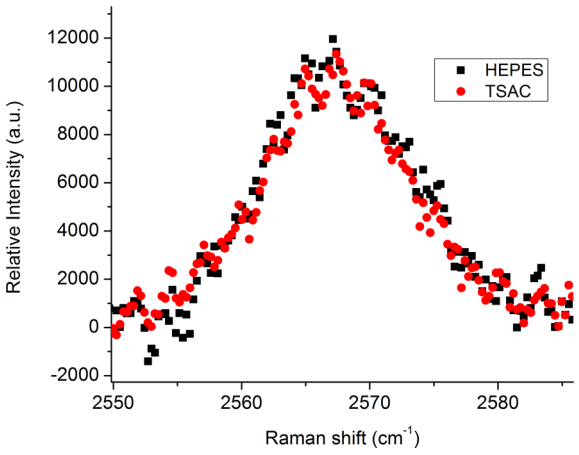


Figure 3. High-resolution Raman S−H stretching bands for rmCK in HEPES buffer and rmCK under TSAC-forming conditions. Each band was fit well to a single Gaussian.

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

(see the Supporting Information). The peak S−H stretching frequency, 2567 cm^{-1} in each spectrum in Figure 3, indicates that the S−H band is not in a fully bulk-like aqueous environment but that the S−H moiety is acting uniformly as a H-bond donor to a weak H-bond acceptor.

The S−H band frequency is mainly a reporter of the participation of the S−H moiety as either a H-bond donor or acceptor. The S−H band's frequency is not strongly sensitive to water or alcohols as H-bond acceptors; the frequency varies significantly only in model compound/solvent studies when the S−H moiety either donates or accepts H-bonds (usually from amides) through direct noncovalent interactions. The S−H line width is much greater in water than in nonpolar, non-H-bonding solvents, likely because water can act as either a weak H-bond donor or an acceptor. Thus, bulk liquid water leads to a larger inhomogeneous distribution of S−H frequencies. Because of its active participation in H-bonding, the S−H moiety is relatively "sticky" to particular parts of its protein environment and could become "stuck" in a particular orientation because of a specific noncovalent attractive interaction. This possible invariance in the SH's interactions with its local environment would be reported by a narrow line width.

The S−H bands in Figure 2 indicate a single dominant environment for the Cys283 thiol in which it acts as a H-bond donor to the same weak acceptor with very little variation in the local structure. The frequency is not low enough to suggest that the acceptor is an amide; the H-bond acceptor in question could be a very weakly accepting amide residue or a temporarily confined, nonbulk water molecule whose identity is not clear from experiment alone.

The invariant S−H band of Cys283 between the apo and TSAC forms indicates that unmodified Cys283 not only is unable to report any possible local differences in the two active sites but also is not sensitive to the presence or absence of the substrate. The S−H frequency indicates that Cys283 forms a weak H-bond, and the relatively narrow line shape (as compared to any of the model compound results in the Supporting Information) suggests that Cys283 preferentially clings to one particular H-bonding partner through all of the conformational changes that the active site undergoes. This specific thiol−acceptor interaction might help to shape and solidify one piece of the active site to best accommodate creatine when it docks, without Cys283 directly interacting with the substrate. If the S−H group were permanently locked into such a H-bond donating conformation, it would not be able to report on other proximate structural changes. This appears to be the case, meaning that this particular S−H vibration is not capable of reporting on any larger structural changes, such as the opening and closing of the active site, because of its homogeneous local interactions.

In both apo and TSAC forms of rmCK, the presence of the Cys283 S−H vibration in the Raman spectrum indicates that Cys283 is clearly in its neutral, protonated form. The S−H stretching band for unmodified, apo rmCK was further observed at several pH values (see the Supporting Information), with its intensity nearly unchanged over all physiologically and biochemically relevant pH values.

MD Simulations of Unmodified rmCK. The relatively short duration of our MD simulations means that the rmCK dimer's simulated structure is not necessarily located at a global thermodynamic minimum but rather remains constrained to a geometry that is very similar to that of the crystal structure used to initiate each simulation. However, we are confident that a well-equilibrated 5 ns production run provides enough sampling to accurately portray all of the local interactions of Cys283, which are what determine the S−H frequency.

MD simulations with protonated Cys283 (Figures 4 and 5) indicate that the thiol of Cys283 spends the majority of its time as a H-bond donor to a single, long-resident water molecule. Figure 4 shows the persistent geometry of Cys283 donating a H-bond to a single persistent water molecule, and Figure 5 shows the time-dependent orientation of the S−H bond through an MD simulation with clear repeating plateaus at the angle associated with formation of this H-bond. The simulated

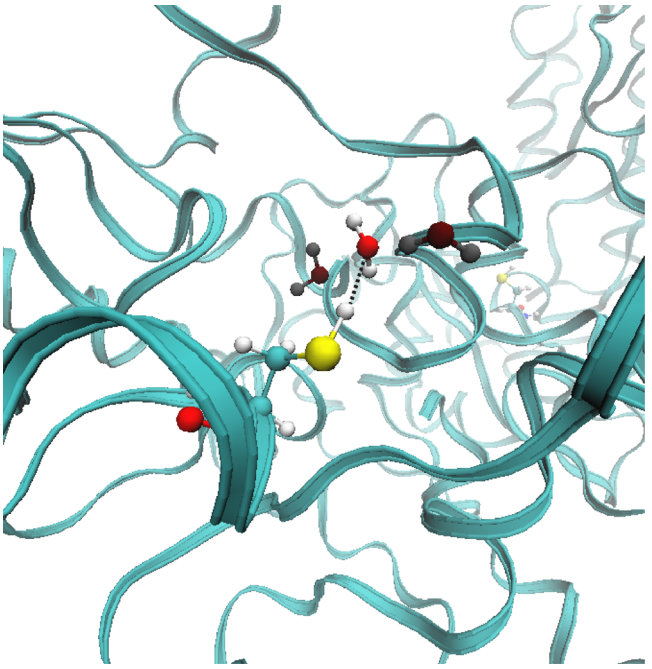


Figure 4. Typical single-frame simulated structure of the rmCK active site in which the Cys283 thiol is constrained by the donation of a hydrogen bond to a single, long-resident water molecule. The Cys283 sulfur atom is colored yellow.

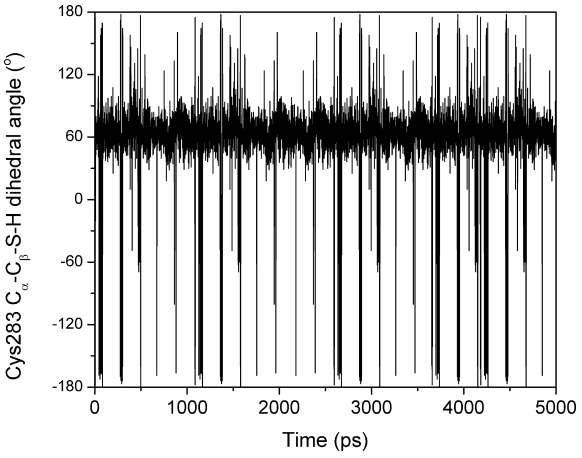


Figure 5. Simulated trajectory for the C_α−C_β−S−H dihedral angle of Cys283 in rmCK, which displays the orientation of the S−H bond with respect to relatively fixed backbone atoms. Many nearly static periods at the dihedral angle of ~65°, which correspond to structures like that shown in in Figure 4, are observed.

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

mean residence time of this water molecule is ~50 ps (see the Supporting Information), during whose residence the Cys283 thiol is confined to one specific orientation. The simulations agree with the proposition that Cys283 is usually H-bonded to a single acceptor, which happens here to be a water molecule.

This observation of the relatively invariant thiol orientation at Cys283 suggests that the catalytically important, but functionally unnecessary, thiol at this site in rmCK organizes the water and H-bonding structure in the active site, along the lines of the structure in Figure 4. It has been known for some time that the thiols of protonated cysteine residues have physical properties that can lead to unique noncovalent interactions with neighboring structural units,[64] with two recent examples that assign an irreplaceable and functionally defining role for a free thiol.[65,66] The active-site thiol is highly conserved across all CK's and in invertebrate AK's; this transient water-structuring role of cysteine is novel and might persist across homologous enzymes from different species.

The invariance of the S−H's simulated geometry and interactions with its local environment agree with the conclusion that this S−H vibrational frequency is not capable of reporting on larger structural changes related to the symmetry or asymmetry of the dimeric enzyme. This very local reporting capacity is likely a general feature of S−H vibrations, whose frequencies are strongly determined by very local interactions with specific H-bond partners.

Yield and Activity of Cyanylated rmCK. $CK(SCN)_2$ was produced in at least 90% yield according to UV−vis spectroscopic monitoring of reaction side products, with the yield calculated on the basis of successful modification of both Cys283 sites or one cysteine per monomer.[5,18] $CK(SCN)_1$ was produced via reaction with NTCB in approximately 70% yield, with a clear plateau before the reaction yield reached a level equivalent to the modification of one thiol per dimer, or 50% of one cysteine per monomer. Reaction with NTCB led to a quantified level of side products consistent with modification of one of every two reactive thiols. Our modification yield for $CK(SCN)_1$ is consistent with that of Degani et al. when using NTCB. This yield is circumstantially consistent with the cyanylation of one reactive thiol on each dimer, and it is not possible that more than one CN group could be added per dimeric enzyme on a sample-wide scale. See the Supporting Information for thiol modification reaction profiles for the two different cyanylation reactions.

The differential reactivity of the two active-site thiols to NTCB and several other reagents suggests that there is an intrinsic asymmetry between the two active sites of the rmCK dimer. Whether that asymmetry exists before the introduction of the reagent, or before the introduction of the substrate to the unmodified enzyme, is an open question that this work intends to answer.

Activity assays for unmodified rmCK and $CK(SCN)_2$ using NADH as a reporter (see the Supporting Information) indicated that $CK(SCN)_2$ catalyzes the phosphorylation of creatine at a rate that is 47% compared to that of the same reaction catalyzed by the unmodified enzyme, compared to the value of 75% previously reported by Degani. The conditions of this assay essentially measure the turnover number, or k_{cat}, of the enzyme. In both our hands and Degani's, the doubly cyanylated enzyme is still active with a moderate deceleration in the enzyme turnover after near-quantitative cyanylation of Cys283 at both active sites. This substantial remaining enzymatic activity indicates that although the active-site geometry has been changed to some extent by the conversion of Cys283 to C*283 (a net change of only one atom), the native function of the enzyme remains. This also means that we can reasonably expect that $CK(SCN)_2$ will be able to form the TSAC because it is an active enzyme.

The NTCB-modified enzyme is not active, an observation consistent with the conclusion of Degani that while half of the CK active sites were cyanylated, the other half were effectively blocked from further reaction.[5,18] Thus, the inactive $CK(SCN)_1$ cannot form a TSAC.

Infrared Spectroscopy of Cyanylated rmCK Variants. The infrared spectra in the CN stretching region for the apo forms of $CK(SCN)_2$ and $CK(SCN)_1$ are shown in Figure 6,

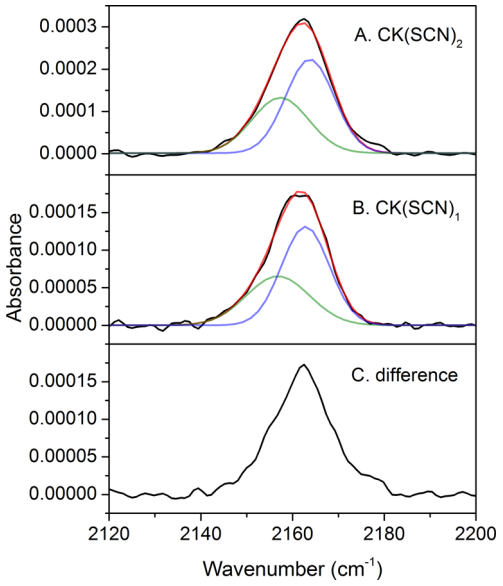


Figure 6. IR CN stretching band of (A) $CK(SCN)_2$ and (B) $CK(SCN)_1$ in 0.5× TAE buffer (pH 7.0) and (C) difference spectrum calculated from $CK(SCN)_2 - 2(CK(SCN)_1)$. Data are colored black. Displayed fits (red) in frames A and B are to two Gaussians at water-exposed (blue) and water-excluded (green) frequencies (see the Supporting Information for quantitative details of fits).

along with two-Gaussian fits to the spectra and a "difference spectrum" calculated by subtracting the spectrum of $CK(SCN)_1$ from the spectrum of $CN(SCN)_2$. The "difference spectrum" visualizes the virtual spectrum of a probe in the alternate active site that has not been converted to C* in $CK(SCN)_1$. All of these spectra are nearly identical within the experimental signal-to-noise level. In both apo samples, $CK(SCN)_1$ and $CN(SCN)_2$, the CN stretching band at C*283 is broad and asymmetric, with a line width greater than any observed for this probe group in all previously reported studies from this laboratory.

The CN stretching bands observed in Figure 6 for the apo-cyanylated rmCK's are atypical of other such bands reported for C* in a number of different protein and peptide systems, including an enzymatic active site,[37,41] a protein−protein interface,[39,40] a disordered protein−protein complex,[44] and a peptide−lipid bilayer interface.[42,45] Specifically, the bands here are notably asymmetric and much broader than for the same side chain in many other proteins and peptides, even in those where the probe is located in a relatively well-structured protein environment. The anomalous width and shape are likely caused by the C*283 infrared probe groups' local environments inside

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

the rmCK active sites. The possibly bisignate line shape is due to some combination of the effects of local electrostatics in the active site,[67] the degree of solvent exposure of the probe group,[35,44] and picosecond time scale local dynamics.[36,68]

Because the line width of the SCN stretching band is typically narrower and more symmetric than any of the cyanylated rmCK spectra, the apo spectra in panels A and B of Figure 6 were each fit to two Gaussian subpopulations of variable frequencies and widths. In each case, the two-Gaussian fit indicates two spectral subpopulations, one at ~2163 cm^{-1} (which is the typical frequency of a completely water-exposed C* residue) and one at ~2157 cm^{-1} (a frequency typically associated with near-complete water exclusion).[35,36,44,45,57] (See the Supporting Information for the quantitative information from these fits.) Alternate single-Gaussian fits to these CN bands did not adequately reproduce the spectral line shapes, and this single-distribution failure strongly suggests that multiple spectral subpopulations are the best explanation for the broadness and asymmetry observed in Figure 6.

A possible interpretation of the two (or at least multiple) spectral subpopulations is that some of the active sites of apo $CK(SCN)_2$ (Figure 6a) are water-filled and some are in a more "closed" conformation that excludes the solvent from the SCN group, and thus, the probe group reports two different solvation environments. Such transient "opening and closing" of the active site was previously postulated for the *Torpedo californica* creatine kinase based on its TSAC crystal structure,[7] and there have been discussions in this field about the possible significance of loop motions (perhaps in the 62−70 loop indicated in Figure 1) that might temporarily "seal" each active site.[69] Because both Cys283 sites are accessible to DTNB and NaCN to become modified, their two environments likely exchange with each other dynamically such that each active site spends part of its time "closed" and the remaining time "open" to solution and the entrance of the substrate. The extinction coefficient of the SCN band has been shown to decrease in nonaqueous environments:[36] a lower spectral intensity in the lower-frequency subpopulation would support the hypothesis that each of Figure 6 apo spectra displays two nearly equal populations, one in which the SCN probe moiety is solvent-exposed and one in which it is water-excluded. However, another possible interpretation of the CN stretching band results is that the artificial probe side chain (which is certainly larger than the native thiol of cysteine) has a limited number of discrete preferred orientations with respect to the surrounding active-site structure, and that directional electrostatic effects and/or differential water exposure of these selected orientations leads to different CN stretching frequencies.

The CN stretching band of the single cyanylated active site in $CK(SCN)_1$ (Figure 6b) is nearly indistinguishable in shape from that of $CK(SCN)_2$. The spectra in panels A and B of Figures 6 were collected at nearly the same rmCK concentration, and the CN bands' relative spectral intensities agree with our conclusion based on the cyanylation reaction yields that one of every two cysteines was cyanylated by NTCB. The calculated difference spectrum (Figure 6c), which is meant to represent the "ghost" probe spectrum for the second active site that is not cyanylated by NTCB, is also nearly identical (within the experimental signal:noise ratio and spectral resolution) to the spectra of $CK(SCN)_2$ and $CK(SCN)_1$. Together, these comparisons suggest that in the apo form of the dimeric enzyme, both active sites can sample all possible active-site geometries to which the vibrational probe group is sensitive. The time scale of exchange between these active-site/probe conformations is presumably slower than the picosecond time scale, but the broad CN line shape precludes any clear conclusions from the one-dimensional IR spectrum about the exchange dynamics between different probe environments. (However, a two-dimensional IR spectrum would be able to place some clear bounds on the exchange time scale between CN frequency subpopulations.[68,70]) The observation of identical, multiple-population probe spectra in both $CK(SCN)_2$ and $CK(SCN)_1$ agrees with the general hypothesis that the rmCK dimer is already dynamically asymmetric in its all-apo form without any conformational bias applied by the substrate.

The infrared spectra of $CK(SCN)_2$ and the TSAC of $CK(SCN)_2$ in HEPES buffer are shown in Figure 7. The substrate-free HEPES spectrum is slightly different from those in Figure 6, likely because the enzyme is not in a completely apo form because of the absence of EDTA from the buffer. There is a very slight shift of the entire spectrum to a lower frequency upon introduction of the substrate and cofactors, but the overall spectrum still has a shape that indicates a range of spectral subpopulations that is quite similar to both the non-TSAC HEPES spectrum and all spectra in Figure 6. The slight red shift of the entire spectrum might be consistent with a small degree of solvent exclusion from the active site introduced by the presence of these new species, but none of the TSAC components appears to interact directly with the SCN probe group. This observation is consistent with the lack of change in the immediate environment of unmodified Cys283 observed above using its S−H stretching band.

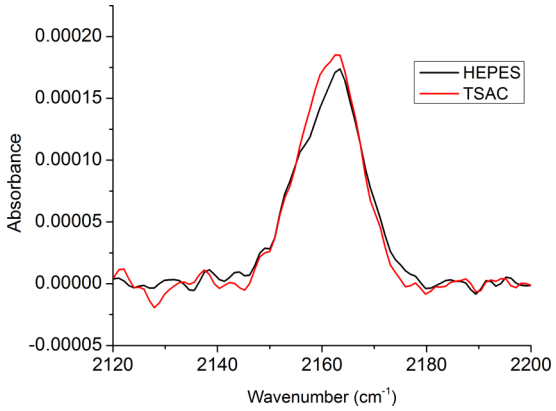


Figure 7. IR CN stretching bands of $CK(SCN)_2$ in HEPES buffer (black) and under TSAC conditions (red).

Molecular dynamics simulations explicitly including the artificial SCN probe group, along either the lines of those recently performed by the Webb group[39,40,71] or those used to simulate vibrational spectra by the Hammes-Schiffer group,[55] could help to clarify the interpretation of the broad and asymmetric CN stretching bands from C*283. The usefulness of such simulations would depend on the fidelity of the force field parameters assumed for the artificial functional group, and any MD simulations on a system as large as CK, which presumably has many possible large-amplitude motions, could suffer from undersampling of the conformational distribution. However, the large range of CN stretching frequencies observed here necessarily comes from a distribution of environments in which the probe group can be both strongly water-exposed and completely water-excluded, so simulations

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

that focus on the SCN group orientation and on the distance and interactions between the SCN group and nearby water molecules should provide a strong interpretive guide for the results in Figures 6 and 7.

Molecular Dynamics Simulations of rmCK with Cyanylated Cys283. Because the orientation of the SCN group is not clear from crystal structures that contain only the unmodified Cys283 thiol group, 12 short simulations were run in which the SCN group was placed in different starting orientations by varying the two degrees of rotational freedom (around the β carbon and around the S atom) in the C* side chain. This strategy is very similar to the "umbrella sampling" approach to simulating C* side chains at known sites in Ras complexes with unknown probe group orientations.[71] In the case of C*283, the C* side chain was held at a fixed orientation throughout the equilibration period, and then the SCN group's orientation was followed during the first few nanoseconds of the production run. The SCN group's orientation as a function of time during the first nanosecond of simulation, starting from 12 different possible orientations in the apo dimer structure, is shown in Figure 8. In these relatively short trajectories, the

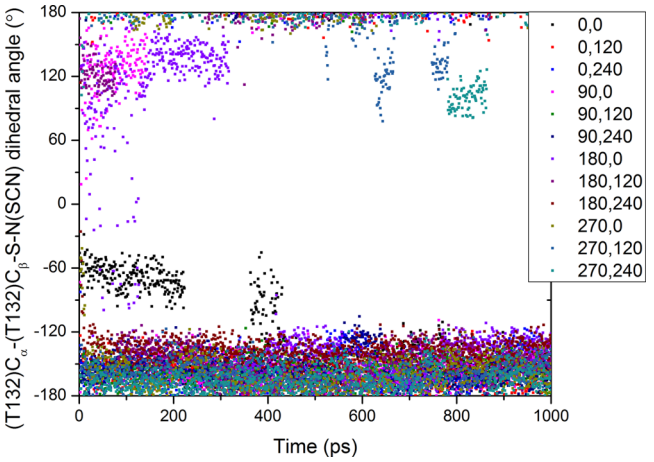


Figure 8. Simulated trajectories for CN-modified rmCK starting from several different starting orientations of the SCN group attached to Cys283. The (T132)C_α–(T132)C_β–S(283)–N(SCN) dihedral angle quantifies the orientation of the SCN group with respect to remote atoms that are relatively fixed throughout this simulation. Each of 12 starting orientations (with rotameric variations at both the C_β and S atoms of the modified Cys283 side chain) quickly collapses to the same, relatively static orientation of the SCN group.

overall structure is expected to remain strongly determined by the starting crystallographic geometry, but the local structure of the SCN group and in its close vicinity is sufficiently well sampled to reach clear conclusions about the SCN group's local environment and any factors that might influence the CN stretching frequency.

From all sampled starting orientations, the SCN group is observed to collapse within picoseconds to a dihedral angle of ~180°, which is the most sterically preferred orientation of this particular alkane-linked function group in which the SCN group points directly away from the backbone with the C* side chain in an extended conformation. The SCN group then remains in that orientation, with only very minor and nonpersistent fluctuations away from it, in all trajectories. The trajectories in Figure 8 suggest that there is one single strongly preferred orientation of the C*283 nitrile group in cyanylated apo rmCK. This result rules out heterogeneous orientations of the probe group in the active site as a possible cause of the broad and asymmetric CN line shapes observed in Figures 6 and 7.

If SCN group local orientations are not a major factor in determining the line shape, then interactions between the SCN group and water are likely the major determinant of the CN stretching frequency. A single representative snapshot that captures the preferred orientation of the C* side chain surrounded by the rest of the active site and its water molecules is shown in Figure 9. The SCN group's orientation is

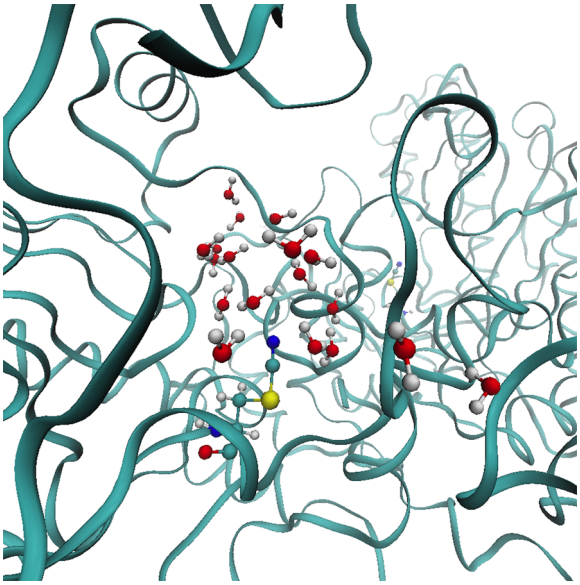


Figure 9. Typical single-frame simulated structure for the CN-modified rmCK active site, with the Cys283-SCN probe group oriented in its relatively static orientation (associated with the persistent geometry in Figure 8). The SCN group is near the bottom of the frame, with the sulfur atom colored yellow. Several water molecules have dynamic access to the N atom of the SCN group.

constrained by steric interactions around the S and C atoms, but the N atom is quite close to several water molecules that can donate weak, transient hydrogen bonds to the N atom. Such weak H-bonding interactions are exactly what cause the blue shift of the nitrile CN stretching frequency in the presence of water molecules.[56] In our MD trajectories, the water molecules near the SCN group and closest to the N atom were observed to fluctuate quickly without any strong orientational preferences or long residence times. This lack of persistent H-bonding is consistent with what was observed in simulations[57] and experimental studies[36] on methyl thiocyanate, in which long-lived 1:1 H-bonded complexes between the SCN group and water molecules were not observed.

There is a relatively large and mobile pool of water molecules close to the N atom of the SCN group in all of our apo simulations, even when the simulations are initiated from the asymmetric dimer TSAC geometry with the substrate and cofactors removed. These MD simulations of cyanylated rmCK's clearly explain the observation of a large CN stretching spectral population centered around 2163 cm^{-1}, and they rule out the local orientation of the SCN group as a strong determinant of the CN frequency but do not provide an explanation for the lower-frequency and apparently water-excluded spectral population that leads to the observed

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83–95

asymmetry of the CN line shape. That population must come from a substantial degree of water exclusion in C*283's vicinity, which could only occur because of some large structural change that closes or otherwise purges water from at least part of the active site.

A circumstantial explanation for this water-excluded probe population might be something in rmCK akin to the loop motion postulated for the *T. californica* CK.[6,7,72] Such loop motion has not been directly implicated in structural studies of rmCK (although it does appear circumstantially in the TSAC structure in Figure 1), but if there were such a large-amplitude motion capable of closing one of the two active sites, its presence in rmCK would fully explain the observed CN stretching spectra for Cys283-cyanylated rmCK's. Observation of such a large-amplitude motion that is not already implicated by existing structures is outside the scope of our current MD simulation capabilities, because this work requires an all-atom, explicit-solvent representation of the relatively large protein to accurately model the SCN probe group and its environment. However, the clearest explanation for both the IR data in Figures 6 and 7 and the differential reactivity of the two active-site thiols is that the two active sites dynamically interchange their structures, with negative cooperativity, between "open" and "closed" conformations regardless of any interaction with substrate.

Dynamic Asymmetry in rmCK. The TSAC crystal structures of several CK's indicate that the homodimeric enzyme is functionally asymmetric, with negative cooperativity between sites, during the execution of its catalytic process. However, the crystal structure of the apo rabbit muscle MM enzyme is symmetric. This has been interpreted as evidence of "induced fit" of the active site by the insertion of the substrate. On the other hand, differential reactivity of the conserved active-site thiol argues toward a conformational selection-type model in which active-site asymmetry between the two subunits is a latent, rather than substrate-induced, phenomenon and the substrate selects and temporarily locks down a particular dimeric conformation. Our solution-phase spectroscopic results at Cys283 provide a methodological bridge between these two prior observations.

The S−H stretching signals for the unmodified Cys283 residues suggest a specific local role for the Cys283 residue that both suggests a novel function for a free cysteine residue and precludes the S−H vibration's ability to report on anything other than its local noncovalent interactions with water. In our analysis, the S−H stretching band of unmodified Cys283 does not report on any larger structural changes related to asymmetry or active-site opening and closing.

The infrared results for the cyanylated active-site cysteines in rmCK disagree with the notion of dimeric asymmetry via induced fit. They suggest strongly that the asymmetry of rmCK is intrinsic to the apo protein in solution and is not only induced by the introduction of substrate. Even the introduction of substrate, and the formation of the asymmetric TSAC, does not lead to substantial changes in the IR spectrum of the active-site vibrational probe groups or the Raman spectrum of the unmodified Cys283 side chains at the same sites. This further suggests that rmCK prefers to be asymmetric in solution regardless of the presence of substrate, with dynamic switching of each active site between different conformations. Almost any possible time scale for the exchange between different active-site and dimer conformations is expected to be substantially slower than the intrinsic time scale of the IR spectroscopic measurements in Figures 6 and 7. Another, slower kinetic technique such as stopped-flow mixing paired with thiol-reactive reagents might be able to indicate the time scale of the exchange between "open" and "closed" active sites, especially if it is milliseconds or slower. Alternatively, an NMR experiment using a ^{13}C-labeled SCN group[35,73,74] could also provide a bound different from vibrational spectroscopy on the exchange process, although NMR could be quite challenging with a dimeric complex of >80 kDa that is expected to display exchange dynamics that could effect the ^{13}C chemical shift of the isotopically labeled probe group. With this in mind, we isolated CK($S^{13}CN$)$_2$ but were unable to observe a ^{13}C signal in the thiocyanate δ^{13}C range (see the Supporting Information), which could indicate indirectly that the active-site structural exchange dynamics fall on an NMR-affecting time scale and thus could greatly broaden the probe's ^{13}C NMR signal. However, if the probe group were strongly oriented in a slowly tumbling complex, as is suggested by our MD results where the SCN group prefers one specific orientation inside the protein, the probe group's ^{13}C NMR band should also be very broad regardless of any exchange dynamics.

Several general questions are raised by this work regarding the importance of asymmetry, complementarity, and dynamics of structural switching in biomolecules. If rmCK is latently asymmetric in a way that appears to be encoded by the interactions between subunits, why should this confer an evolutionary advantage? Could this possibly confer tighter regulatory capability and make CK more useful as an on-demand cellular agent? How far does this behavior stretch: is it limited to rmCK, is it perpetuated throughout CK's and perhaps across other phosphagen kinases, and does it extend to other enzymatic systems that function as oligomers with apparent half-of-sites or partial-sites reactivity? And does rmCK, as Goodsell suggests,[3] provide a somewhat primordial example of a functional and encoded asymmetry that may be used in biological assemblies from dimeric enzymes to motor proteins and viral capsids with many copies of the same modular subunit in complementary structural and functional roles? To answer each of these questions, techniques that intrinsically incorporate dynamics of many time scales (such as vibrational spectroscopy, but also including NMR and other *in vitro* approaches) will need to be developed and carefully utilized to elucidate any underlying principles of dynamic biomolecular design that explicitly, or implicitly, encode dynamic asymmetry in their functional conformations.

Vibrational Spectroscopy Methodology. This is the first study in which S−H bands from free thiols and CN stretching bands from SCN labels at the same sites are used in a complementary fashion to provide a comprehensive viewpoint of the local environment around the cysteine residue of interest. The two approaches provide very different information about the same site, partly because the S−H group interacts so directly with its local environment and the SCN group does not. In this case, the S−H band reports directly on the surprising long-lived "attachment" of the free thiol at Cys283 to one particular water molecule and suggests a new possible role for the thiol in CK's catalytic reaction. The S−H band also provides a signal that directly reports the protonation state of the thiol group. The SCN band, on the other hand, appears to originate from a residue that points uniformly in one direction and thus it is able to report on more global changes to the shape of the active site. This general approach, using both the free thiol's S−H band and the modified SCN group, could

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

provide a robust strategy for determining the role of free cysteines, which appear infrequently in proteins but whose biochemical roles appear to be much more varied than the possible formation of disulfide bonds or other redox-affected species. These vibrational spectroscopy techniques directly provide information about the distribution of environments around both unlabeled and labeled natively occurring cysteines. The use of molecular dynamics simulations to interpret the vibrational spectroscopy results augments these experiments' utility in determining the exact biochemical participation of cysteine residues.

CONCLUSIONS

The Raman S−H stretching vibrations of the catalytically important Cys283 residue in rmCK display a single narrow band that persists at all functionally relevant pH's and appears at a frequency that suggests the donation of a hydrogen bond from the thiol to a homogeneous, weak H-bond acceptor. MD simulations reveal that H-bond acceptor to be a single, long-lived water molecule that resides for tens of picoseconds and pins the thiol group into a single orientation. This observation suggests that the novel functional role of Cys283 is to organize the H-bonding network in its end of the active site through interaction with its long-lived partner water molecule.

The CN stretching vibrations of rmCK's that are cyanylated at Cys283 display a broad and asymmetric band that suggests at least two spectral subpopulations. MD simulations rule out the local orientation of the SCN labeling group as a major factor in the CN line shape, leaving the overall asymmetry of the two active sites as the best explanation for the observed label infrared spectra. Because the broad spectrum is invariant to whether one or both Cys283's are labeled, the asymmetry of the two active sites must be dynamic and complementary. The dynamic asymmetry suggested by this work contradicts the substrate-induced fit explanation used to explain asymmetric TSAC crystal structures of several CK's, and it raises a number of questions about the possible functional importance of dynamic asymmetry in multimeric protein complexes.

This study also presents the first example of the complementary use of S−H vibrations of free cysteines and CN vibrations of cyanylated versions of the same side chains to report on complementary features of the same protein sites. In this case, and in general, the S−H mode reports mainly on the local interactions of the thiol with its environment, while the CN group is a more passive observer of its location and thus can report on slightly larger-scale phenomena like local folding or the opening and closing of an active site. This could be a generally useful strategy for determining the roles of free cysteines in proteins where the cysteines have clear functional importance.

ASSOCIATED CONTENT

Supporting Information

Detailed experimental methods, quantitative details of fits to spectral data, and complementary experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: clonderg@haverford.edu.

Funding

D.M.S. acknowledges summer support from a medical education grant to Haverford College by the Howard Hughes Medical Institute. K.W.H. acknowledges a Beckman Scholarship from the Arnold and Mabel Beckman Foundation. This work was begun through funding from Haverford College and a New Faculty Start-Up Award to C.H.L. from the Dreyfus Foundation and is partly based on work currently funded by National Science Foundation Grant CHE-1150727.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Haverford College colleague Karin Akerfeldt is thanked for stimulating discussions, and Heather A. McMahon and Stephen Griffiths are acknowledged for experimental efforts associated with this enzyme but not included here.

ABBREVIATIONS

AK, arginine kinase; CK, creatine kinase; rmCK, rabbit muscle creatine kinase; hmCK, human muscle creatine kinase; TSAC, transition state analogue complex; pCR, phosphocreatine; ADP, adenosine diphosphate; DTNB, Ellman's reagent or 5,5′-dithiobis(2-nitrobenzoic acid); TNB, 2-nitrobenzoic acid-5-thiolate; NTCB, 2-nitro-5-thiocyanobenzoic acid; TAE, Tris-acetate EDTA; C*, cyanylated cysteine or β-thiocyanoalanine.

REFERENCES

(1) Noda, L. H., Kuby, S. A., and Lardy, H. A. (1954) Adenosinetriphosphate(ATP)-creatine transphosphorylase. II. Homogeneity and physicochemical properties. *J. Biol. Chem. 209*, 203−210.

(2) Kuby, S. A., Noda, L. H., and Lardy, H. A. (1954) Adenosinetriphosphate-creatine transphosphorylase. I. Isolation of the crystalline enzyme from rabbit muscle. *J. Biol. Chem. 209*, 191−201.

(3) Goodsell, D. S., and Olson, A. J. (2000) Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct. 29*, 105−153.

(4) Bickerstaff, G. F., and Price, N. C. (1978) Creatine kinase: Review of some recent work on mechanism and subunit behavior of enzyme. *Int. J. Biochem. 9*, 1−8.

(5) Degani, Y., and Degani, C. (1980) Enzymes with asymmetrically arranged subunits. *Trends Biochem. Sci. 5*, 337−341.

(6) McLeish, M. J., and Kenyon, G. L. (2005) Relating structure to mechanism in creatine kinase. *Crit. Rev. Biochem. Mol. Biol. 40*, 1−20.

(7) Lahiri, S. D., Wang, P. F., Babbitt, P. C., McLeish, M. J., Kenyon, G. L., and Allen, K. N. (2002) The 2.1 Å structure of *Torpedo californica* creatine kinase complexed with the ADP-Mg^{2+}-NO^{3-}-creatine transition-state analogue complex. *Biochemistry 41*, 13861−13867.

(8) Ohren, J. F., Kundracik, M. L., Borders, C. L., Edmiston, P., and Viola, R. E. (2007) Structural asymmetry and intersubunit communication in muscle creatine kinase. *Acta Crystallogr. D63*, 381−389.

(9) Bong, S. M., Moon, J. H., Nam, K. H., Lee, K. S., Chi, Y. M., and Hwang, K. Y. (2008) Structural studies of human brain-type creatine kinase complexed with the ADP-Mg^{2+}-NO^{3-}-creatine transition-state analogue complex. *FEBS Lett. 582*, 3959−3965.

(10) Levitzki, A., Stallcup, W. B., and Koshland, D. E. (1971) Half-of-the-sites reactivity and the conformational states of cytidine triphosphate synthetase. *Biochemistry 10*, 3371−3378.

(11) Kaldis, P., and Wallimann, T. (1995) Functional differences between dimeric and octameric mitochondrial creatine kinase. *Biochem. J. 308*, 623−627.

(12) Rao, J. K., Bujacz, G., and Wlodawer, A. (1998) Crystal structure of rabbit muscle creatine kinase. *FEBS Lett. 439*, 133−137.

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

(13) Eder, M., Schlattner, U., Becker, A., Wallimann, T., Kabsch, W., and Fritz-Wolf, K. (1999) Crystal structure of brain-type creatine kinase at 1.41 angstrom resolution. *Protein Sci. 8*, 2258−2269.
(14) Eder, M., Fritz-Wolf, K., Kabsch, W., Wallimann, T., and Schlattner, U. (2000) Crystal structure of human ubiquitous mitochondrial creatine kinase. *Proteins: Struct., Funct., Genet. 39*, 216−225.
(15) Liu, N., Wang, J. S., Wang, W. D., and Pan, J. C. (2011) The role of Cys271 in conformational changes of arginine kinase. *Int. J. Biol. Macromol. 49*, 98−102.
(16) Hornemann, T., Rutishauser, D., and Wallimann, T. (2000) Why is creatine kinase a dimer? Evidence for cooperativity between the two subunits. *Biochim. Biophys. Acta 1480*, 365−373.
(17) Degani, C., and Degani, Y. (1980) Further evidence of nonsymmetric subunit association and intersubunit cooperativity in creatine kinase: Subunit-selective modifications by 2,4-dinitrophenylthiocyanate. *J. Biol. Chem. 255*, 8221−8228.
(18) Degani, Y., and Degani, C. (1979) Subunit-selective chemical modifications of creatine kinase: Evidence for asymmetrical association of the subunits. *Biochemistry 18*, 5917−5923.
(19) Maggio, E. T., Kenyon, G. L., Markham, G. D., and Reed, G. H. (1977) Properties of a CH_3-blocked creatine kinase with altered catalytic activity: Kinetic consequences of presence of blocking group. *J. Biol. Chem. 252*, 1202−1207.
(20) Markham, G. D., Reed, G. H., Maggio, E. T., and Kenyon, G. L. (1977) Magnetic resonance studies of 3 forms of creatine kinase: Comparison of properties of native, CH_3S-blocked and H_2NCOCH_2-blocked enzymes. *J. Biol. Chem. 252*, 1197−1201.
(21) Nevinsky, G. A., Ankilova, V. N., Lavrik, O. I., Mkrtchyan, Z. S., Nersesova, L. S., and Akopyan, J. I. (1982) Functional non-identity of creatine kinase subunits of rabbit skeletal muscle. *FEBS Lett. 149*, 36−40.
(22) Price, N. C., and Hunter, M. G. (1976) Non-identical behavior of subunits of rabbit muscle creatine kinase. *Biochim. Biophys. Acta 445*, 364−376.
(23) Furter, R., Furtergraves, E. M., and Wallimann, T. (1993) Creatine kinase: The reactive cysteine is required for synergism but is nonessential for catalysis. *Biochemistry 32*, 7022−7029.
(24) Wang, P. F., McLeish, M. J., Kneen, M. M., Lee, G., and Kenyon, G. L. (2001) An unusually low pK_a for Cys282 in the active site of human muscle creatine kinase. *Biochemistry 40*, 11698−11705.
(25) Wang, P. F., Flynn, A. J., Naor, M. M., Jensen, J. H., Cui, G. L., Merz, K. M., Kenyon, G. L., and McLeish, M. J. (2006) Exploring the role of the active site cysteine in human muscle creatine kinase. *Biochemistry 45*, 11464−11472.
(26) Cox, J. M., Chan, C. A., Chan, C., Jourden, M. J., Jorjorian, A. D., Brym, M. J., Snider, M. J., Borders, C. L., and Edmiston, P. L. (2003) Generation of an active monomer of rabbit muscle creatine kinase by site-directed mutagenesis: The effect of quaternary structure on catalysis and stability. *Biochemistry 42*, 1863−1871.
(27) Awama, A. M., Mazon, H., Vial, C., and Marcillat, O. (2007) Despite its high similarity with monomeric arginine kinase, muscle creatine kinase is only enzymatically active as a dimer. *Arch. Biochem. Biophys. 458*, 158−166.
(28) Carey, P. R. (1999) Raman spectroscopy, the sleeping giant in structural biology, awakes. *J. Biol. Chem. 274*, 26625−26628.
(29) Lindquist, B. A., Furse, K. E., and Corcelli, S. A. (2009) Nitrile groups as vibrational probes of biomolecular structure and dynamics: An overview. *Phys. Chem. Chem. Phys. 11*, 8119−8132.
(30) Waegele, M. M., Culik, R. M., and Gai, F. (2011) Site-Specific Spectroscopic Reporters of the Local Electric Field, Hydration, Structure, and Dynamics of Biomolecules. *J. Phys. Chem. Lett. 2*, 2598−2609.
(31) Alben, J. O., Bare, G. H., and Bromberg, P. A. (1974) Sulfhydryl groups as a new molecular probe at α1-β1 interface in hemoglobin using Fourier-transform infrared spectroscopy. *Nature 252*, 736−738.
(32) Dong, A. C., Huang, P., Caughey, B., and Caughey, W. S. (1995) Infrared analysis of ligand-induced and oxidation-induced conformational changes in hemoglobins and myoglobins. *Arch. Biochem. Biophys. 316*, 893−898.
(33) Tuma, R., Vohnik, S., Li, H. M., and Thomas, G. J. (1993) Cysteine conformation and sulfhydryl interactions in proteins and viruses. 3. Quantitative measurements of the Raman S-H band intensity and frequency. *Biophys. J. 65*, 1066−1072.
(34) Li, H. M., and Thomas, G. J. (1991) Cysteine conformation and sulfhydryl interactions in proteins and viruses. 1. Correlation of the Raman S-H band with hydrogen-bonding and intramolecular geometry in model compounds. *J. Am. Chem. Soc. 113*, 456−462.
(35) Fafarman, A. T., Sigala, P. A., Herschlag, D., and Boxer, S. G. (2010) Decomposition of Vibrational Shifts of Nitriles into Electrostatic and Hydrogen-Bonding Effects. *J. Am. Chem. Soc. 132*, 12811−12813.
(36) Maienschein-Cline, M. G., and Londergan, C. H. (2007) The CN stretching band of aliphatic thiocyanate is sensitive to solvent dynamics and specific solvation. *J. Phys. Chem. A 111*, 10020−10025.
(37) Fafarman, A. T., Sigala, P. A., Schwans, J. P., Fenn, T. D., Herschlag, D., and Boxer, S. G. (2012) Quantitative, directional measurement of electric field heterogeneity in the active site of ketosteroid isomerase. *Proc. Natl. Acad. Sci. U.S.A. 109*, E299−E308.
(38) Fafarman, A. T., Webb, L. J., Chuang, J. I., and Boxer, S. G. (2006) Site-specific conversion of cysteine thiols into thiocyanate creates an IR probe for electric fields in proteins. *J. Am. Chem. Soc. 128*, 13356−13357.
(39) Stafford, A. J., Ensign, D. L., and Webb, L. J. (2010) Vibrational Stark Effect Spectroscopy at the Interface of Ras and Rap1A Bound to the Ras Binding Domain of RalGDS Reveals an Electrostatic Mechanism for Protein-Protein Interaction. *J. Phys. Chem. B 114*, 15331−15344.
(40) Stafford, A. J., Walker, D. M., and Webb, L. J. (2012) Electrostatic Effects of Mutations of Ras Glutamine 61 Measured Using Vibrational Spectroscopy of a Thiocyanate Probe. *Biochemistry 51*, 2757−2767.
(41) Sigala, P. A., Fafarman, A. T., Bogard, P. E., Boxer, S. G., and Herschlag, D. (2007) Do ligand binding and solvent exclusion alter the electrostatic character within the oxyanion hole of an enzymatic active site? *J. Am. Chem. Soc. 129*, 12104−12105.
(42) Alfieri, K. N., Vienneau, A. R., and Londergan, C. H. (2011) Using Infrared Spectroscopy of Cyanylated Cysteine to Map Membrane Binding Structure and Orientation of the Hybrid Antimicrobial Peptide CM15. *Biochemistry 50*, 11097−11108.
(43) Edelstein, L., Stetz, M. A., McMahon, H. A., and Londergan, C. H. (2010) The Effects of α-Helical Structure and Cyanylated Cysteine on Each Other. *J. Phys. Chem. B 114*, 4931−4936.
(44) Bischak, C. G., Longhi, S., Snead, D. M., Constanzo, S., Terrer, E., and Londergan, C. H. (2010) Probing structural transitions in the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by vibrational spectroscopy of cyanylated cysteines. *Biophys. J. 99*, 1676−1683.
(45) McMahon, H. A., Alfieri, K. N., Clark, C. A. A., and Londergan, C. H. (2010) Cyanylated Cysteine: A Covalently Attached Vibrational Probe of Protein-Lipid Contacts. *J. Phys. Chem. Lett. 1*, 850−855.
(46) Hoffman, K. W., Romei, M. G., and Londergan, C. H. (2013) A New Raman Spectroscopic Probe of Both the Protonation State and Noncovalent Interactions of Histidine Residues. *J. Phys. Chem. A 117*, 5987−5996.
(47) Degani, Y., Neumann, H., and Patchorn, A. (1970) Selective cyanylation of sulfhydryl groups. *J. Am. Chem. Soc. 92*, 6969−6970.
(48) Degani, Y., and Patchorn, A. (1974) Cyanylation of sulfhydryl groups by 2-nitro-5-thiobenzoic acid: High-yeld modification and cleavage of peptides and cysteine residues. *Biochemistry 13*, 1−11.
(49) Riddles, P. W., Blakeley, R. L., and Zerner, B. (1979) Ellman's reagent: 5,5′-Dithiobis(2-nitrobenzoic acid)—a re-examination. *Anal. Biochem. 94*, 75−81.
(50) Riddles, P. W., Blakeley, R. L., and Zerner, B. (1983) Reassessment of Ellman reagent. *Methods Enzymol. 91*, 49−60.
(51) Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel,

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95

D., Hess, B., and Lindahl, E. (2013) GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics 29*, 845−854.
(52) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605−1612.
(53) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf. 65*, 712−725.
(54) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparuson of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(55) Layfield, J. P., and Hammes-Schiffer, S. (2013) Calculation of Vibrational Shifts of Nitrile Probes in the Active Site of Ketosteroid Isomerase upon Ligand Binding. *J. Am. Chem. Soc. 135*, 717−725.
(56) Choi, J. H., Oh, K. I., Lee, H., Lee, C., and Cho, M. (2008) Nitrile and thiocyanate IR probes: Quantum chemistry calculation studies and multivariate least-square fitting analysis. *J. Chem. Phys. 128*, 134506.
(57) Oh, K. I., Choi, J. H., Lee, J. H., Han, J. B., Lee, H., and Cho, M. (2008) Nitrile and thiocyanate IR probes: Molecular dynamics simulation studies. *J. Chem. Phys. 128*, 154504.
(58) Lee, C. T., Yang, W. T., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functionla of the electron-density. *Phys. Rev. B 37*, 785−789.
(59) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Rob, M. A., Cheeseman, J. R., Montgomery, J. A., Jr., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., et al. (2003) *Gaussian 03*, Gaussian, Inc., Wallingford, CT.
(60) Raso, S. W., Clark, P. L., Haase-Pettingell, C., King, J., and Thomas, G. J. (2001) Distinct cysteine sulfhydryl environments detected by analysis of Raman S-H markers of Cys → Ser mutant proteins. *J. Mol. Biol. 307*, 899−911.
(61) Rodriguez-Casado, A., Moore, S. D., Prevelige, P. E., and Thomas, G. J. (2001) Structure of bacteriophage P22 portal protein in relation to assembly: Investigation by Raman spectroscopy. *Biochemistry 40*, 13583−13591.
(62) Rodriguez-Casado, A., and Thomas, G. J. (2003) Structural roles of subunit cysteines in the folding and assembly of the DNA packaging machine (portal) of bacteriophage P22. *Biochemistry 42*, 3437−3445.
(63) Li, H. M., Hanson, C., Fuchs, J. A., Woodward, C., and Thomas, G. J. (1993) Determination of the pKa values of active-center cysteines, cysteines-32 and cysteines-35, in *Escherichia coli* thioredoxin by Raman spectroscopy. *Biochemistry 32*, 5800−5808.
(64) Gregoret, L. M., Rader, S. D., Fletterick, R. J., and Cohen, F. E. (1991) Hydrogen bonds involving sulfur atoms in proteins. *Proteins 9*, 99−107.
(65) Olivella, M., Caltabiano, G., and Cordomi, A. (2013) The role of cysteine 6.47 in class A GPCRs. *BMC Struct. Biol. 13*, 10.
(66) Ozcan, A., Olmez, E. O., and Alakent, B. (2013) Effects of protonation state of Asp181 and position of active site water molecules on the conformation of PTP1B. *Proteins: Struct., Funct., Bioinf. 81*, 788−804.
(67) Bagchi, S., Fried, S. D., and Boxer, S. G. (2012) A Solvatochromic Model Calibrates Nitriles' Vibrational Frequencies to Electrostatic Fields. *J. Am. Chem. Soc. 134*, 10373−10376.
(68) Bagchi, S., Boxer, S. G., and Fayer, M. D. (2012) Ribonuclease S Dynamics Measured Using a Nitrile Label with 2D IR Vibrational Echo Spectroscopy. *J. Phys. Chem. B 116*, 4034−4042.
(69) Kenyon, G. L., Wang, P. F., Flynn, A. J., McLeish, M. J., Lahiri, S. D., and Allen, K. N. (2003) Loop movements and catalysis in creatine kinase. *Biochemistry 42*, 8621.
(70) Kim, Y. S., and Hochstrasser, R. M. (2005) Chemical exchange 2D IR of hydrogen-bond making and breaking. *Proc. Natl. Acad. Sci. U.S.A. 102*, 11185−11190.
(71) Walker, D. M., Hayes, E. C., and Webb, L. J. (2013) Vibrational Stark effect spectroscopy reveals complementary electrostatic fields created by protein-protein binding at the interface of Ras and Ral. *Phys. Chem. Chem. Phys. 15*, 12241−12252.
(72) Wang, P. F., Flynn, A. J., McLeish, M. J., and Kenyon, G. L. (2005) Loop movement and catalysis in creatine kinase. *IUBMB Life 57*, 355−362.
(73) Doherty, G. M., Motherway, R., Mayhew, S. G., and Malthouse, J. P. (1992) ^{13}C NMR of cyanylated flavodoxin from *Megasphaera elsdenii* and of thiocyanate model compounds. *Biochemistry 31*, 7922−7930.
(74) Waldman, A. D. B., Birdsall, B., Roberts, G. C. K., and Holbrook, J. J. (1986) ^{13}C-NMR and transient kinetic studies on lactate-dehydrogenase Cys(CN)165-^{13}C: Direct measurement of a rate-limiting rearrangements in protein structure. *Biochim. Biophys. Acta 870*, 102−111.

DOI: 10.1021/bi5008063
Biochemistry 2015, 54, 83−95