BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 4987–4997 4987
DOI: 10.1021/bi1001745

Alkyl Isocyanides Serve as Transition State Analogues for Ligand Entry and Exit in Myoglobin[†]

George C. Blouin,[‡] Rachel L. Schweers,[§] and John S. Olson*

Department of Biochemistry and Cell Biology and W. M. Keck Center for Computational Biology, Rice University, Houston, Texas 77005 [‡]*Present address: Physics Laboratory, Biophysics Group, National Institute of Standards and Technology, Gaithersburg, MD* [§]*Present address: Katten Muchin Rosenman LLP, Chicago, IL*

Received February 4, 2010; Revised Manuscript Received May 13, 2010

ABSTRACT: Alkyl isocyanides (CNRs) identify pathways for diatomic ligand movement into and out of Mb, with their side chains acting as transition state analogues. The bound alkyl groups point either into the back of the distal pocket (*in* conformation, $\nu_{CN} \approx 2070-2090$ cm^{-1}), which allows hydrogen bond donation from His64(E7) to the isocyano group, or toward solvent through an open His(E7) channel (*out* conformation, $\nu_{CN} \approx 2110-2130$ cm^{-1}), which prevents polar interactions with the isocyano atoms. Fractions of the *in* conformer (F_{in}) were measured by FTIR spectroscopy for methyl through *n*-pentyl isocyanide bound to a series of 20 different distal pocket mutants of sperm whale myoglobin and found to be governed by the ease of rotation of the His(E7) side chain, distal pocket volume and steric interactions, and, for the longer isocyanides, the unfavorable hydrophobic effect of placing their terminal carbon atoms into the solvent phase in the *out* conformation. There are strong correlations between the fraction of *in* conformer, F_{in}, for long-chain MbCNR complexes measured by FTIR spectroscopy, the fraction of geminate recombination of photodissociated O_2, and the bimolecular rates of O_2 entry into the distal pocket. These correlations indicate that alkyl isocyanides serve as transition state analogues for the movement of O_2 into and out of the binding pocket of Mb.

In the preceding paper, we used crystal structures of alkyl isocyanides bound to Mb[1] (MbCNRs) to visualize the entry and exit channels involved in diatomic ligand binding (*1*). Isocyanides coordinate to the Mb heme iron, and the alkyl side chain (R group) is relatively free to rotate and occupy the lowest free energy positions adjacent to the active site, which are presumably the spaces and channels most accessible to the movement of the physiologically relevant ligands, O_2, CO, and NO. Two major isocyanide conformations are observed in the crystal structures of native and wild-type (wt) MbCNR complexes that correlate with the two major isocyanide stretching frequency bands observed in solution FTIR spectra (*1*, *2*). The low frequency (~2075 cm^{-1}) ν_{CN} band represents the *in* conformer, in which the bound alkyl chain is pointing toward the protein interior. In this orientation, the distal histidine gate is closed and forms a hydrogen bond to the bound isocyano group and decreases its bond order. The high-frequency (~2125 cm^{-1}) ν_{CN} band represents the *out* conformer, in which the ligand side chain is pointing toward solvent through an open E7 channel created by outward movement of the His64 side chain. In this orientation, the bound isocyano group is in an apolar environment, which favors a larger bond order (*2*).

In the case of bound *n*-butyl isocyanide (CNC4), the fractions of *in* and *out* conformations are roughly equal in solution, indicating that the unfavorable free energy for packing the four butyl carbon atoms in the interior of the distal pocket is roughly equal to that for opening the distal histidine gate and filling the E7 channel (Figure 1). In the previous paper we suggested that the inward- and outward-pointing positions of the butyl side chain represent respectively the initial and final pathways for the movement of photodissociated and thermally dissociated diatomic ligands.

Olson, Gibson, Phillips, and co-workers (*3*, *4*) have proposed that diatomic ligands enter mammalian Mbs and hemoglobins through the distal His(E7) gate and are then captured in the interior portion of the distal pocket and adjacent Xe binding cavities before binding to the iron atom. They suggested the metaphor of catching of a ball in a baseball glove. Opening of the glove by movement of the thumb is analogous to the outward movement of the His(E7) side chain, and catching in the web of the glove is analogous to ligand capture in the interior cavities of the globin.

Ligand dissociation represents a reversal of this process and can be examined directly by laser photolysis of Mb−ligand complexes. Gibson, Scott, and Olson (*3*, *4*) described the internal ligand movements and geminate recombination reaction in terms of a side path scheme. Photolysis breaks the Fe−ligand bond causing the dissociated ligand to move into the initial B-state, which is located directly above the interior pyrrole rings of the heme group. From there, the ligand can either move into the Xe4

[†] Supported by U.S. Public Health Service Grants GM 35649 (J.S.O.) and HL 47020 (J.S.O.) and Grant C-612 (J.S.O.) from the Robert A. Welch Foundation. R.L.S. and G.C.B. were recipients of traineeships from The Houston Area Molecular Biophysics Predoctroral Training Grant GM08280.

*To whom correspondence should be addressed at the Department of Biochemistry and Cell Biology, Rice University. Telephone: 713-348-4762. Fax: 713-348-5154. E-mail: olson@rice.edu.

[1] Abbreviations: Mb, myoglobin; wt, wild type; CNR, alkyl isocyanide, where R = C1, C2, C3, C4, C5, and C6 for methyl, ethyl, *n*-propyl, *n*-butyl, *n*-pentyl, and *n*-hexyl groups; FTIR, Fourier transform infrared spectroscopy; ν_{CN}, isocyano group stretching frequency; F_{in}, fraction of CNRs that point into the Mb binding pocket as measured by FTIR; k', association rate; k, dissociation rate; K_a, equilibrium association constant calculated as k'/k; F_{gem}, fraction of geminate recombination.

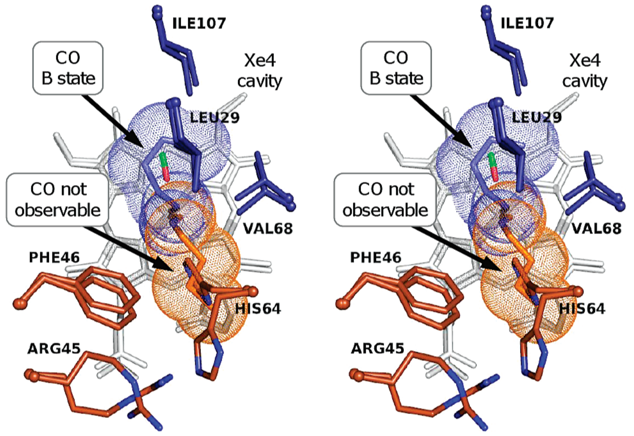


FIGURE 1: Stereo view of amino acids selected for mutation near bound CNRs. The binding pockets are shown for native (pH 7) and wt (pH 9) MbCNC4 (104m and 111m, respectively (*1*)) and are globally aligned by their Cα atoms. Mutations of the residues shown in dark blue are expected to affect the free energy of CNC4 in the *in* conformation (slate blue) by increased or decreased steric hindrance, whereas those in brown should selectively affect the *out* conformer of CNC4 (orange). The heme groups are in white, and van der Waals spheres for CNC4 are shown to provide an indication of residue/ligand packing. The alkyl groups of bound CNRs may act as transitions state analogues for CO in the photodissociated B state or during entry into the binding pocket through the His64 "gate" (marked as CO not observable). CO in the dissociated B state has been observed directly in low-temperature and time-resolved crystallography experiments, whereas, to date, ligand movement through the E7 gate has not been observed directly.

and Xe1 sites (states C and D), rebind, or escape through the His(E7) gate (Figure 1). Movement into the C and D sites represents a side path from which the ligand rarely leaves the protein. Instead, it returns to the distal pocket to either rebind to the iron (geminate recombination) or escape from the protein.

We hypothesize that the CNR alkyl C atoms in the *in* conformation simulate the initial trajectory taken by photodissociated diatomic ligands into the B-state site and that the alkyl carbon atoms in the *out* conformation mimic O_2, CO, and NO at points along their escape path through the E7 channel (Figure 1 (*4*, *5*)). To examine the correspondence between the side path scheme and the isocyanide conformation, we systematically measured the effects of key distal pocket mutations at positions 29, 45, 46, 68, and 107 (Figure 1) on the fraction of *in* versus *out* CNR conformers as measured by FTIR spectroscopy. The results were then compared to previously measured rates of O_2 entry and fractions of geminate rebinding for the same library of Mb variants (*4*). The resultant correlations between these independently measured parameters are striking and support strongly the idea that the positions of bound isocyanide side chains do identify the pathways for diatomic ligand movement in myoglobin.

MATERIALS AND METHODS

Detailed methods can be found in the first paper of this series (*2*). Recombinant Mbs were expressed in *Escherichia coli* and purified as described by Springer and Sligar (*6*) and modified by Carver et al. (*7*). The CNRs were synthesized by our group or by Mark Hargrove's group at Iowa State University using the methods of Casanova et al. (*8*−*10*).

Geminate recombination measurements were carried out for MbCNRs in N_2-equilibrated, pH 7.0 buffer containing 100 mM potassium phosphate, 1 mM EDTA, 50−100 μM Mb, and 100−1000 μM CNR. Rate constants for CNR binding were measured as described in ref *11*. We chose not to determine all of the rate parameters for the isocyanides because our goal was to correlate the position of the ligand side chain with the rate constants for diatomic ligand binding and not to study isocyanide binding *per se* as has been done previously for various naturally occurring globins (*10*, *12*) and His64 mutants (*11*).

FTIR measurements were conducted by loading into a 40 μm CaF_2 cell an ∼20 μL sample of 2−5 mM Mb and 1× and 5× molar equivalents of CNR and sodium dithionite, respectively. All solutions contained N_2-equilibrated, 100 mM potassium phosphate, 1 mM EDTA, pH 7.0 buffer. The FTIR spectra were collected on a Nicolet Nexus 470 FTIR spectrometer and have a 1 or 2 cm^{-1} resolution.

RESULTS

Effects of Mutations That Alter the Flexibility of the His64(E7) Side Chain. Substitutions of His64 clearly alter ligand entry and exit rates. Scott et al. (*4*) observed a roughly linear dependence of the O_2 entry rate on the size of the side chain at position E7 for the series of relatively apolar amino acids, Gly, Ala, Val, Leu, Phe, and Trp. They also observed that mutations at locations spatially near the distal histidine, Arg45(CD3) and Phe46(CD4), could cause significant increases in the rates of ligand entry and exit by enhancing the mobility of the His64 side chain.

The Arg45(CD3) guanidinium group is part of a hydrogen-bonding network that includes the Nδ of His64, the heme propionates, and water molecules when His64 is in the closed conformation (*13*, *14*). Disruption of this hydrogen-bonding network moderately enhances the rates of ligand entry and exit by destabilizing the closed conformation of the histidine gate. The Phe46(CD4) benzyl group sterically hinders outward rotation of the His64 side chain. In the crystal structure of F46V MbCO, His64 rotates upward to fill the space left by removal of the phenyl group and, in doing so, opens a pore between the binding pocket and the solvent phase near the heme propionates (*15*). Smith et al. (*1*) observed an identical opening of the His64 side chain in the F46V MbCNC4 crystal structure. In the latter case, the pore is filled with the butyl group of the bound isocyanide in the *out* conformation (Figure 2C).

The FTIR measurements shown in Figure 2 support the idea that the R45K, R45E, and F46V mutations allow a greater ease of opening the E7 gate. Although the conservative R45K substitution retains a positively charged amino acid at position 45(CD3), it causes small increases in the amplitudes of the ν_{CN} *out* peaks in all of the R45K MbCNR spectra (Figure 2A). Loss of the three N atoms of the arginine guanidinium partially disrupts the hydrogen-bonding network that stabilizes His64 in the closed conformation and allows more of the bound ligands to point outward. The larger increase in the amount of *out* conformer for the R45E mutant is due both to disruption of the hydrogen bond network and to introduction of a negative charge near His(E7). The net change in charge of −2 due to the R45E mutation increases the pK_a of the His64 side chain, its extent of protonation at pH 7, and the amount of outward rotation of the resulting imidazolium cation. Carver et al. (*13*) studied a similar mutant, K45E pig Mb, and found that CNC3 and CNC4 had 10-fold increases in affinity over the wt protein (Table 1, last column). This effect is most likely caused by selective relief of steric hindrance on the bound CNR in the *out* conformation.

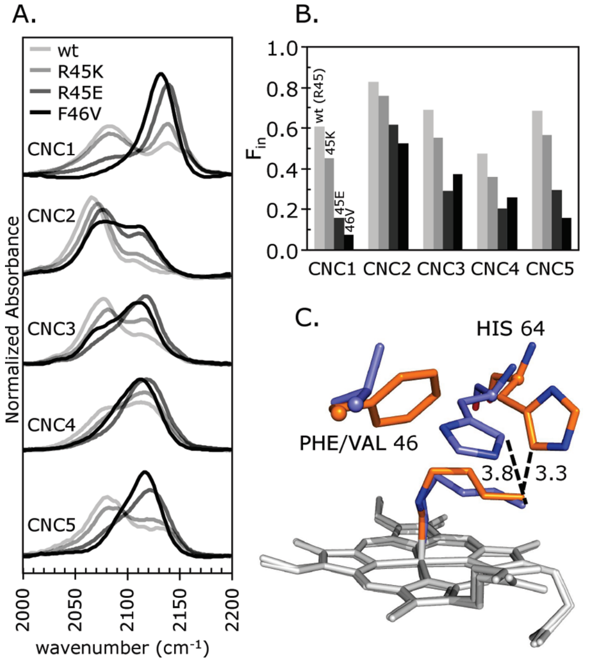


FIGURE 2: Effect of mutating residues Arg45 and Phe46. (A) The ν_{CN} peaks in the FTIR spectra of methyl through pentyl isocyanides (CNC1−CNC5) bound to wt, R45K, R45E, and F46V Mbs. (B) The fraction of *in* conformers (F_{in}) for each mutant in panel A as a function of CNR size. These values indicate that the distal histidine is pushed outward more easily in the order wt < R45K < R45E ≈ F46V by the alkyl tail of each ligand. (C) Structures of native (orange) and F46V (slate blue) MbCNC4 (104m and 101m, respectively (*1*)), globally aligned by their Cα atoms. The smaller Val46 side chain allows the His64 imidazole to rotate away from the terminal carbon of the ligand and relieve steric hindrance in the *out* conformation.

The infrared spectra for methyl through pentyl isocyanide bound to F46V Mb are also shown in Figure 2A. The fractions of *in* conformers for F46V Mb are roughly 2−10-fold lower than their corresponding wt Mb values (F_{in}, Figure 2B). Clearly, the histidine gate in this mutant is easily opened (Figure 2C). Although the ligand binding data are limited to CNC4, the *out* conformer in the F46V mutant is less hindered than either conformer in wt Mb, and as a result, the affinity of F46V Mb for CNC4 is roughly 10-fold higher than that for wt Mb and similar to that for K45E pig Mb (Table 1).

For wt Mb and the R45K and R45E mutants, a plot of F_{in} versus CNR size has an undulating up−down−up shape with CNC2 and CNC4 at the vertices (Figure 2B, tops of the bars for each variant). Blouin and Olson (*2*) postulated that the increase in F_{in} for CNC2 compared to CNC1 is due to the inductive effect of the ethyl group, which favors more unpaired electron density on the N atom for interaction with His64. The space occupied by the ligand Cβ is unhindered and is in the same location as the initial position of photodissociated ligands (Figure 1). The decrease in F_{in} for CNC2−CNC4 is due to steric constraints on the Cγ and Cδ atoms of the alkyl side chain, which begin to push the *n*-propyl and *n*-butyl groups out of the pocket and into the E7 channel. CNC5 is long enough to access solvent directly outside the binding pocket in the *out* conformer, where it disrupts the water hydrogen-bonding lattice. Thus, for the longest ligands, the hydrophobic effect becomes dominant, and the fraction *in* becomes higher again for both CNC5 (Figure 2) and CNC6 (see Figure 6C).

However, for F46V Mb, F_{in} continues to decrease for CNC5, indicating that despite the unfavorable hydrophobic effect the open conformation is more favorable than packing the five alkyl carbon atoms in the interior of the distal pocket. Superimposition

Table 1: FTIR Spectroscopy and Ligand Binding Kinetic Parameters Measured for MbCNR Complexes with R45 Mutations E and K and F46V[a]

MbCNR complex	FTIR spectral data			CNR binding data			
	in ν_{CN} (cm^{-1})	*out* ν_{CN} (cm^{-1})	F_{in}	F_{gem}	k' (μM^{-1} s^{-1})	k (s^{-1})	K_a (μM^{-1})
wt							
CNC1	2083	2139	0.60	0.80	0.12[b]	4.3[b]	0.028[b]
CNC2	2065	2106	0.82	0.95	0.074[b]	0.27[b]	0.27[b]
CNC3	2077	2112	0.69	0.79	0.043[b]	0.33[b]	0.13[b]
CNC4	2083	2113	0.47	0.63	0.029[b]	0.60[b]	0.048[b]
CNC5	2081	2131	0.68	0.89	0.030[c]	0.44[c]	0.069[c]
R45E							
CNC1	2083	2139	0.16	0.59[d]	0.23[d]	4.7[d]	0.049[d]
CNC2	2077	2112	0.61		0.29[d]	1.0[d]	0.28[d]
CNC3	2077	2118	0.29		0.23[d]	0.46[d]	0.51[d]
CNC4	2077	2118	0.20		0.18[d]	0.36[d]	0.51[d]
CNC5	2081	2121	0.29				
R45K							
CNC1	2082	2139	0.45	0.78[d]	0.11[d]	4.4[d]	0.025[d]
CNC2	2071	2107	0.76		0.098[d]	0.63[d]	0.15[d]
CNC3	2082	2116	0.55		0.063[d]	0.70[d]	0.090[d]
CNC4	2078	2116	0.36		0.040[d]	0.50[d]	0.081[d]
CNC5	2085	2122	0.57				
F46V							
CNC1	2083	2131	0.07	0.47			
CNC2	2079	2110	0.52	0.54			
CNC3	2075	2110	0.37	0.42			
CNC4	2074	2112	0.26	0.30	0.22[e]	0.63[e]	0.36[e]
CNC5	2074	2118	0.16				

[a] The parameter abbreviations are described in footnote [1]. The *in* and *out* ν_{CN} values are the low- and high-frequency isocyano FTIR peaks, respectively. Blank cells indicate parameters that have not been measured. [b] From ref *11*. [c] Values are for native sperm whale Mb and are averages from refs *10* and *12*. [d] Values are for K45E and wt pig Mb from ref *13*. [e] From ref *16*.

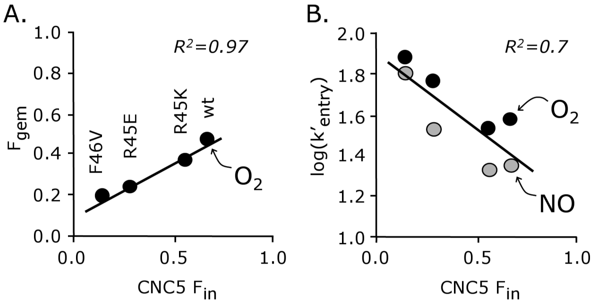


FIGURE 3: The CNR *out* conformation acts as a transition state analogue for diatomic ligand entry. (A) Correlations between the fraction of geminate recombination of photolyzed O_2, F_{gem,O_2}, and the fraction of bound CNC5 that adopts the *in* conformation, $F_{in,CNC5}$, for Mb mutants with substitutions near the His(E7) gate, including R45K, R45E, and F46V. F_{gem,O_2} values were taken from ref *4*. (B) Correlations for wt, R45K, R45E, and F46V Mbs between the logs of k'_{entry,O_2} or k'_{NO} and $F_{in,CNC5}$.

of the native MbCNC4 ($P2_1$, pH 7) and F46V MbCNC4 ($P6$, pH 9) crystal structures (*1*) shows that, although the distal histidine in both structures is rotated into an open conformation, the space left by the removal of the Phe46 phenyl group allows the His64 imidazole to rotate further away from the ligand alkyl atoms in the *out* conformation. As a result, there is a clearance of 3.8 Å between the closest atom of the His64 side chain (Cδ) and the terminal carbon of the CNC4 alkyl group in F46V Mb. In contrast, this distance is 3.3 Å in the native Mb structure (Figure 2C).

Correlations between F_{in} and Rates of O_2 and NO Entry for Mutations at the E7 Gate. The values of F_{in} for the R45E, R45K, and F46V MbCNC5 complexes show the greatest variation (Figure 2B, last set of bars) and were compared to previously determined rate constants for O_2 entry, k'_{entry,O_2}, and the fraction of O_2 geminate recombination, F_{gem,O_2}, for the same set of Mb variants (Figure 3). As shown in Figure 3, there is a strong linear correlation (R^2 values ≥ 0.9) between F_{in} for CNC5 measured by FTIR and F_{gem,O_2} (*4*). The positive slope confirms that larger F_{in} values, which indicate a higher free energy barrier to E7 gate opening, are associated with larger F_{gem} values for O_2, which indicate greater trapping of diatomic ligands in the distal pocket after photolysis.

The conclusion that His(E7) in Mb controls ligand access to the Mb distal pocket is supported by the inverse correlation of the fraction of *in* CNC5 conformer with both the computed rates of ligand entry for O_2 binding, k'_{entry}, and the directly measured association rate constants for NO binding, k'_{NO}, for the position 45 and 46 mutants (Figure 3B). For NO, the fraction of internal binding versus escape is always ≥ 0.99 due to the high reactivity of the NO radical, and thus the rate-limiting step for NO binding is entry into the distal pocket (*4*). This inverse correlation demonstrates that a lower barrier to opening the histidine gate, indicated by a smaller $F_{in,CNC5}$, predicts significantly higher bimolecular rates of diatomic ligand entry.

Similar, though less strong, correlations of F_{in} for the CNC1−CNC4 complexes of these mutants with F_{gem,O_2}, k'_{entry,O_2}, and k'_{NO} are observed (data not shown). The close proximity of His64 to the Mb heme iron allows any bound CNR, regardless of its size, to act as a probe of the stability of its closed versus open conformations. Longer CNRs are required to study the effects of mutations at positions deeper in the binding pocket and to indicate the importance of capture volume in the reactions of diatomic ligands with Mb.

Effects of Val68(E11) Mutations: Decreasing and Enhancing Direct Steric Hindrance and Blocking Access to the Xe4 Pocket. The distal valine(E11) has multiple effects on ligand binding to Mb including (a) sterically hindering access to the heme iron (*17−19*), (b) affecting the position of the distal histidine in its closed conformation, (c) lining a portion of the E7 channel, and (d) being part of the short path between the initial geminate or B-state and the more interior Xe4 cavity (Figure 1). Thus, mutations at the E11 position have complex, sometimes competing effects on bimolecular and geminate rebinding and on the orientation of bound isocyanides.

The V68A mutation in Mb results in a large decrease in steric pressure on all bound CNRs, as judged by the 10−100-fold increases in overall affinity (K_a values in Table 2). However, the effects of this release of hindrance are different for the short versus the long straight-chained ligands (see the spectra in Figure 4A and the first two bars of each set in Figure 4B). The F_{in} values for CNC1, CNC2, and CNC3 bound to the V68A mutant are all smaller than those for wt Mb, indicating that opening up of the E7 channel due to loss of the adjacent valine $\gamma 2$ methyl group is the dominant effect. In contrast, the F_{in} values for CNC4 and CNC5 bound to V68A Mb are larger than the wt values (Figure 4B, Table 2), indicating that loss of the $\gamma 1$ methyl group creates more space in the back of the distal pocket for a direct connection to the Xe4 cavity and accommodation of the terminal carbon atoms of the *n*-butyl and *n*-pentyl side chains (Figure 1). This increase in space allows the hydrophobic effect to push the side chains of the longer ligands back into the protein interior (Figure 4B, Table 2). These differences roughly correlate with the differences in the fractions of geminate recombination observed for the rebinding of these ligands to wt and V68A Mb after laser photolysis. The short-chain isocyanides show smaller while the longer ligands show larger F_{gem} and F_{in} values (Figure 4B, Table 2).

In contrast to the V68A mutation, which reduces steric hindrance of the bound ligands, the V68I mutation increases direct hindrance significantly and decreases the affinities for O_2, CO, and the alkyl isocyanides CNC1−CNC4 by 4−20-fold (Table 2 (*17*)). The Cδ methyl group of isoleucine is in van der Waals contact with bound CO in the X-ray structure of V68I MbCO and is found in two rotomers indicating significant hindrance and disorder in the liganded complex (Figure 4C, green sticks (*18*)).

The IR spectra of V68I MbCNRs are complex with two to three broad peaks, reflecting the conformational heterogeneity observed in the V68I MbCO crystal structure. The three ν_{CN} bands are assignable based on their alignments with the wt MbCNR peaks. One low-frequency peak probably represents an *in* conformation, and the two high-frequency peaks probably represent *out* conformations. The splitting of the *out* conformer peak may be due to the presence of both Ile68 rotomers shown in Figure 4C. Despite this complexity, it is clear from the spectra in Figure 4A that the dominant effect of the V68I mutation is to push the alkyl isocyanides into an *out* conformation, which opens the His(E7) gate. This effect is most dramatic for the longer isocyanides where $\geq 60\%$ and 100% of CNC4 and CNC5 conformers, respectively, are pointing outward.

The V68F mutation offers a test of the ability of longer alkyl isocyanides to access the distal Xe4 cavity. In the crystal structures of V68F Mb, the benzene ring of Phe68 partially fills and blocks access to the Xe4 cavity (Figure 4D, green sticks). The lack of an effect of the Phe68 mutation on F_{in} for CNC1, CNC2,

Table 2: FTIR Spectroscopy and Ligand Binding Kinetic Parameters Measured for MbCNR Complexes with V68 Mutations A, I, F, L, and T[a]

	FTIR spectral data			CNR binding data			
MbCNR complex	*in* ν_{CN} (cm^{-1})	*out* ν_{CN} (cm^{-1})	F_{in}	F_{gem}	k' (μM^{-1} s^{-1})	k (s^{-1})	K_a (μM^{-1})
wt							
CNC1	2083	2139	0.60	0.80	0.12[b]	4.3[b]	0.028[b]
CNC2	2065	2106	0.82	0.95	0.074[b]	0.27[b]	0.27[b]
CNC3	2077	2112	0.69	0.79	0.043[b]	0.33[b]	0.13[b]
CNC4	2083	2113	0.47	0.63	0.029[b]	0.60[b]	0.048[b]
CNC5	2081	2131	0.68	0.89	0.030[c]	0.44[c]	0.069[c]
CNC6	2078	2137	0.90	0.97	0.037[c]	0.15[c]	0.25[c]
V68A							
CNC1	2100	2137	0.38	0.66	0.38[d]	0.76[d]	0.50[d]
CNC2	2064	2106	0.69	0.83	0.18[d]	0.07[d]	2.6[d]
CNC3	2089	2121	0.61	0.97	0.11[d]	0.022[d]	5.0[d]
CNC4	2079	2121	0.60	0.99	0.14[d]	0.011[d]	13[d]
CNC5	2065	2121	0.75				
V68I							
CNC1	2091	2138	0.21	0.57[e]	0.05[d]	21[d]	0.0024[d]
CNC2	2065	2095	0.53	0.63[e]	0.047[d]	3.4[d]	0.014[d]
CNC3	2073	2116	0.37		0.016[d]	1.5[d]	0.011[d]
CNC4	2092	2115	0.39		0.018[d]	1.5[d]	0.012[d]
CNC5	ND	2118	0.00				
V68F							
CNC1	2098	2139	0.59	0.99	0.013[d]	0.03[d]	0.43[d]
CNC2	2064	2100	0.86	0.99	0.0061[d]	0.0035[d]	1.7[d]
CNC3	2079	2108	0.71	0.96	0.004[d]	0.0019[d]	2.1[d]
CNC4	2079	2108	0.53	0.95	0.0058[d]	0.0077[d]	0.75[d]
CNC5	2079	2107	0.46	0.44			
CNC6	2079	2107	0.40	0.72			
V68L							
CNC1		2140	~0.0		0.36	1.7	0.21
CNC2	2065	2100	~0.8		0.19	0.55	0.35
CNC3		2090	NA		0.26	0.34	0.76
CNC4		2090	NA		0.22	0.13	1.7
CNC5		2100	NA				
V68T							
CNC1	2118	2139	0.64	0.63[f]	0.039[f]	1.2[f]	0.033[f]
CNC2	2085	2109	0.65	0.67[f]	0.018[f]	0.092[f]	0.20[f]
CNC3	2095	2121	0.64		0.0095[f]	0.082[f]	0.12[f]
CNC4	2080	2127	0.67		0.0088[f]	0.16[f]	0.056[f]
CNC5	2074	2142	0.86				

[a] Details regarding the parameters are presented in Table 1. NA = not applicable. ND = not detected. Blank cells indicate parameters that have not been measured. [b] From ref *11*. [c] Values are for native sperm whale Mb and are averages from refs *10* and *12*. [d] From ref *17*. [e] From ref *20*. [f] For V68T pig Mb from ref *21*.

and CNC3 demonstrates that, as in the crystal structure, the 68(E11) phenyl group does not significantly intrude into the distal pocket and hinder bound ligands (Figure 4B, Table 2). F_{in} for CNC4 is also unchanged by this mutation, although in the crystal structure, the Phe68 side chain sterically hinders the terminal carbon of the CNC4 butyl group (Figure 4D). However, the V68F mutation does significantly decrease F_{in} for CNC5 from 0.68 in wt Mb to 0.46 in the mutant (Figure 4B, Table 2). The increase in steric hindrance on CNC5 in the *in* conformation due to blockage of the Xe4 cavity by Phe68 partially overcomes the unfavorable hydrophobic forces on the ligand in the *out* conformation (*2*). The effect of decreasing F_{in} by blocking the Xe4 cavity is even larger for the longer CNC6 ligand, as shown in Figure 6.

The IR spectra of the V68L MbCNR complexes are difficult to interpret because, except for CNC1, a single broad band is observed with a peak maximum midway between that for the normal *in* and *out* peaks of the wt complexes (Table 2; spectra not shown). Consequently, F_{in} values could not be determined or used in the more global analyses described in the Discussion (Table 2). In the crystal structure of V68L metMb, deoxyMb, and MbCO, the isobutyl side chain occupies multiple conformations, which alternatively fill the back of the distal pocket or the space directly above the iron atom, displacing distal pocket water from the unliganded protein (*18*). As a result, an ensemble of isocyanide conformers is most likely present that cannot be described in terms of only one *in* and one *out* orientation.

The effect of increasing the distal pocket polarity was also examined by replacing Val68 with Thr. Following the work of Smerdon et al. (*21*), Li et al. found the V68T mutation causes a 17 cm^{-1} blue shift of the major ν_{CO} band of pig and sperm whale MbCO and attributed this effect to the proximity of the negative dipole of the 68T hydroxyl group to the bound ligand atoms (*22*). There are similar large blue shifts for the low-frequency *in* peaks of CNC1, CNC2, and CNC3 bound to V68T Mb but not for CNC4 and CNC5 (Table 2; spectra not shown).

Effects of Leu29(B10) Mutations. In globins, the B-helix forms part of the roof of the distal binding pocket above the heme, and the B10 side chain extends down toward the bound ligand (Figures 1 and 5C). In mammalian Mbs and Hbs, the native Leu(B10) directs the movement of dissociated ligands

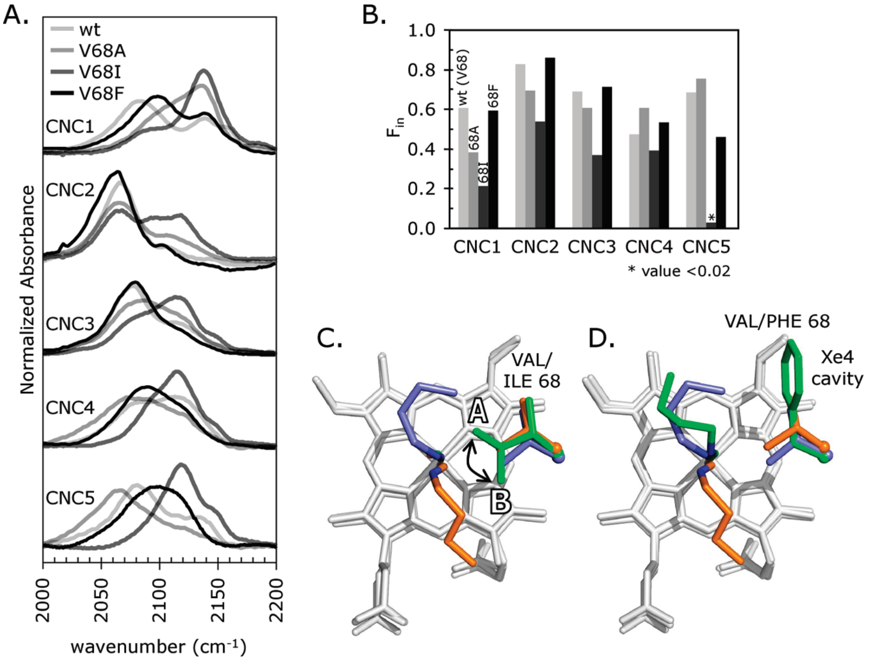


FIGURE 4: Effect of mutating residue V68. (A) FTIR spectra of methyl through pentyl isocyanides (CNC1−CNC5) bound to wt and V68A, V68F, and V68I Mbs. (B) F_{in} for each mutant−isocyanide pair as a function of ligand size. (C) Native (orange) and wt (slate blue) MbCNC4 and V68I MbCO (green) globally aligned by their Cα atoms (PDB IDs 104m and 111m (*1*) and 1mlm (*18*), respectively). The V68I MbCO structure has two Ile68 χ_2 rotomers (A and B). Hemes are in white. (D) The same structural overlay as in panel C, except V68I MbCO is replaced by V68F MbCNC4 (PDB ID 107m (*1*)).

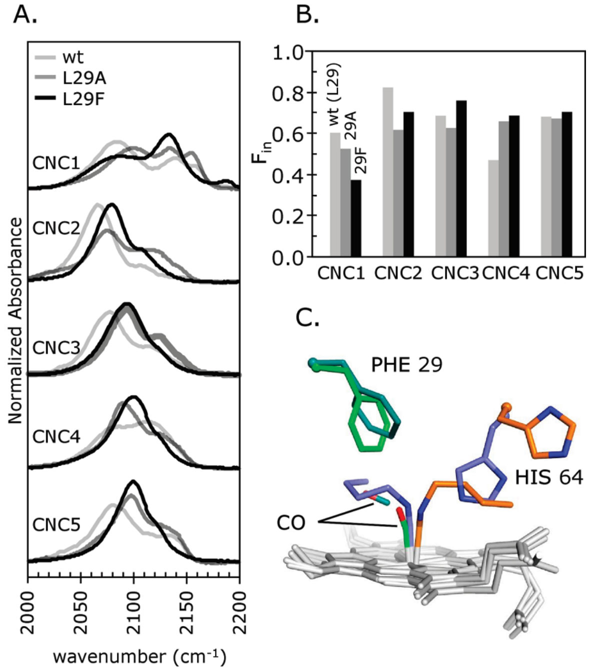


FIGURE 5: Effect of mutating residue Leu29. (A) FTIR spectra of methyl through pentyl isocyanides (CNC1−CNC5) bound to wt, L29A, and L29F Mbs. (B) F_{in} for each mutant−isocyanide pair as a function of ligand size. (C) Structures of native (orange) and wt (slate blue) MbCNC4 and L29F MbCO (light green) and L29F Mb:CO photoproduct (dark green), globally aligned by their Cα atoms (PDB IDs 104m and 111m (*1*) and 2g0r and 2g0v (*23*), respectively; hemes in light gray). The benzyl side chain at residue 29 rotates to accommodate photodissociated CO in the B-state. A similar rotation may explain how the large CNR alkyl groups are accommodated within the L29F Mb binding pocket.

either into the back of the distal pocket or out through the E7 gate and, if increased in size to Phe or Trp, can directly interact with bound diatomic ligands (*7*, *23*−*25*). Substitution of Leu(B10) with the smaller Ala residue significantly increases the binding pocket volume available to bound isocyanides, causing the marked increases in ligand affinities, particularly for the larger isocyanides (Table 3). In contrast, the Phe(B10) replacement causes decreases in isocyanide affinity. However, remarkably, the F_{in} values for L29A and L29F MbCNR complexes determined by ν_{CN} measurements are similar to each other and to those for the wt MbCNRs.

The FTIR results in Table 3 and Figure 5 show that an increase or a decrease in the size of the B10 amino acid causes only small changes in F_{in}, even though ligand affinities change dramatically (Figure 5B, Table 3). The *in* ν_{CN} peak is shifted to higher frequencies for both the L29A and L29F mutants, suggesting a more upright conformation in both cases, and the amount of *in* conformer for the larger CNC4 ligand is significantly greater than for wt Mb. The latter result is easy to understand for the L29A mutant in which the internal distal cavity is much larger than that in wt Mb; however, the increase in F_{in} for Phe(B10) Mb is more difficult to rationalize. Aranda et al. (*23*) have shown that Phe(B10) can move away from photodissociated CO in time-resolved X-ray crystallographic studies of L29F MbCO and does not appear to obstruct the B-state site, which is readily occupied by photodissociated CO (Figure 5C). The FTIR data also indicate that the long alkyl side chains of bound isocyanides are accommodated in the interior distal pocket of L29F MbCNR.

Effects of Ile107(G8) Mutations and Blocking Access to the Xe4 Cavity for Bound CNC5 and CNC6. The Ile107(G8) side chain is located at the back of the distal pocket and serves as a partial barrier to ligand movement into the Xe4 pocket (*4*). Substitution of alanine for isoleucine at G8 enhances the stability of ethyl through pentyl isocyanides in the *in* conformation, with CNC4 giving the greatest F_{in} increase from 0.47 in wt Mb to 0.80 in I107A Mb (Figure 6A, Table 4; spectra not shown). Inspection of the wt MbCNC4 crystal structure shows that the Ile107 Cδ

Table 3: FTIR Spectroscopy and Ligand Binding Kinetic Parameters Measured for MbCNR Complexes with L29 Mutations A and F[a]

MbCNR complex	FTIR spectral data			CNR binding data			
	in ν_{CN} (cm^{-1})	*out* ν_{CN} (cm^{-1})	F_{in}	F_{gem}	k' (μM^{-1} s^{-1})	k (s^{-1})	K_a (μM^{-1})
wt							
CNC1	2083	2139	0.60	0.80	0.12[b]	4.3[b]	0.028[b]
CNC2	2065	2106	0.82	0.95	0.074[b]	0.27[b]	0.27[b]
CNC3	2077	2112	0.69	0.79	0.043[b]	0.33[b]	0.13[b]
CNC4	2083	2113	0.47	0.63	0.029[b]	0.60[b]	0.048[b]
CNC5	2081	2131	0.68	0.89	0.030[c]	0.44[c]	0.069[c]
L29A							
CNC1	2098	2154	0.53	0.62	0.58[d]	7.7[d]	0.075[d]
CNC2	2075	2119	0.62	0.97	0.70[d]	0.63[d]	1.1[d]
CNC3	2093	2123	0.63	0.97	0.45	0.023	20
CNC4	2089	2126	0.66	0.98	2.0	0.020	100
CNC5	2098	2134	0.68				
L29F							
CNC1	2084	2133	0.38	0.29	0.19[d]	57[d]	0.0033[d]
CNC2	2079	2108	0.71	0.44	0.090[d]	4.2[d]	0.021[d]
CNC3	2093	2127	0.76	0.42			
CNC4	2099	2123	0.69	0.28			
CNC5	2098	2124	0.71				

[a] Details regarding the parameters are presented in Table 1. Blank cells indicate parameters that have not been measured. [b] From ref *11*. [c] Values are for native sperm whale Mb and are averages from refs *10* and *12*. [d] From ref *7*.

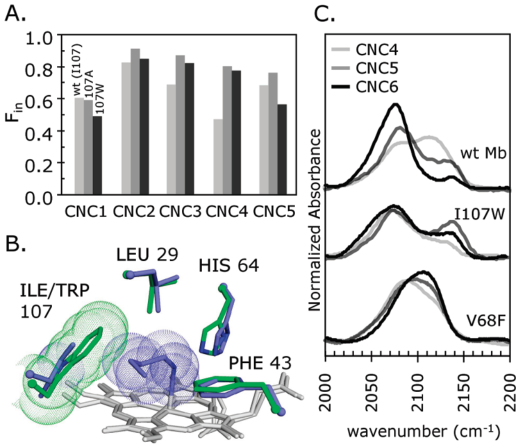


FIGURE 6: Effect of side chain size at residue 107 and blocking the Xe4 pocket. (A) F_{in} plotted against ligand size for methyl through pentyl isocyanides (CNC1−CNC5) bound to wt, I107A, and I107W Mbs. (B) Native MbCNC4 (slate blue; pH 9 structure; PDB ID 105m (*1*)) and I107W Mb (green; ferric form; PDB ID 2ohb (*5*)) globally aligned by their Cα atoms, with van der Waals spheres shown for CNC4 and the Trp107 side chain. The F_{in} value for the actual I107W MbCNC4 complex is unexpectedly high (panel A). (C) FTIR spectra of CNC4−CNC6 bound to wt, I107W, and V68F Mbs. Mutations that block access to the Xe4 site inhibit inward movement of longer CNRs, as shown by increases in the high-frequency *out* conformation band at ~2125 cm^{-1}.

atom contacts the C3 and C4 atoms of the ligand butyl group along its C1−C4 path from the B-state near the heme iron to the Xe4 pocket (Figure 1). The structural and FTIR data indicate that crowding by the Ile107 side chain significantly contributes to the large free energy penalty of adding a fourth carbon to the CNR *n*-alkyl group in the *in* conformation and that the further addition of a fifth carbon is energetically neutral because it is located within the Xe4 cavity.

Surprisingly, the I107W mutation appears to expand the distal pocket. The Trp107 mutant more readily accommodates moderately sized isocyanides, CNC2−CNC4, as judged by significantly larger F_{in} values compared to those for wt Mb (Figure 6A, Table 4). This result is consistent with the 30-fold greater K_a and the larger fraction of geminate recombination for CNC4 binding to I107W Mb than to wt Mb. The structure of I107W metMb shows that the Trp107 indole is sterically restricted by the edge of the heme group from rotating into the B-state associated with the initial CO photoproduct (Figure 6B). However, this mutation does significantly lower the F_{in} values for bound CNC5 and CNC6 compared to their much higher wild-type values (Figure 6A, Table 4). As is the case for V68F Mb (Figure 4B,D), the indole ring of Trp107 appears to block access to the Xe4 cavity and sterically hinders the ends of the longer CNRs as they wrap around the back of the distal pocket.

Complexes of *n*-hexyl isocyanide, CNC6, bound to wt Mb, V68F Mb, and I107W Mb serve as tests of the ease of movement into the Xe4 cavity, which is described as state C in most photodissociation schemes. Figure 6C shows comparisons of the ν_{CN} spectra for these complexes and those for CNC4 and CNC5 bound to the same proteins. As described in the first paper of this series (*2*), there is a progressive increase in the inward-pointing population for CNC4, CNC5, and CNC6 bound to wt Mb, with F_{in} values of 0.47, 0.68, and 0.90, respectively (Figure 6C, top set of spectra; Table 4). The size-dependent increase of the *in* conformation for these three CNRs is due to the terminal alkyl carbon atoms being driven out of the solvent phase and into the protein interior by a hydrophobic effect. There appear to be only small increases in steric hindrance for the C5 and C6 carbon atoms in the *in* conformation because they can be accommodated within the empty Xe4 cavity (*2*). This interpretation is supported by the effects of placing a tryptophan at 107(G8) or phenylalanine at 68(E11), which block access to the Xe4 cavity and partially reverse the trend to higher F_{in} values for the CNC4 to CNC6 series (Figure 6C, Table 4). Thus, when access to the protein interior is sterically obstructed, there is a roughly 40% decrease in the inward-pointing population of the long-chain isocyanides. Under these conditions, the hydrophobic effect driving inward movement roughly equals the steric hindrance caused by forcing the terminal carbon atoms into the protein interior.

Table 4: FTIR Spectroscopy and Ligand Binding Kinetic Parameters Measured for MbCNR Complexes with I107 Mutations A and W[a]

MbCNR complex	FTIR spectral data			CNR binding data			
	in ν_{CN} (cm^{-1})	*out* ν_{CN} (cm^{-1})	F_{in}	F_{gem}	k' (μM^{-1} s^{-1})	k (s^{-1})	K_a (μM^{-1})
wt							
CNC1	2083	2139	0.60	0.80	0.12[b]	4.3[b]	0.028[b]
CNC2	2065	2106	0.82	0.95	0.074[b]	0.27[b]	0.27[b]
CNC3	2077	2112	0.69	0.79	0.043[b]	0.33[b]	0.13[b]
CNC4	2083	2113	0.47	0.63	0.029[b]	0.60[b]	0.048[b]
CNC5	2081	2131	0.68	0.89	0.030[c]	0.44[c]	0.069[c]
CNC6	2078	2137	0.90	0.97	0.037[c]	0.15[c]	0.25[c]
I107A							
CNC1	2089	2141	0.59	0.73	0.091	2.3	0.040
CNC2	2067	2147	0.91	0.92	0.12	0.15	0.80
CNC3	2083	2141	0.87	0.99	0.13	0.025	5.2
CNC4	2081	2143	0.80	0.99	0.19	0.0088	22
CNC5	2079	2144	0.76	0.99			
I107W							
CNC1	2096	2143	0.49	0.87	0.18	0.21	0.86
CNC2	2066	2147	0.85	0.94	0.16	0.025	6.4
CNC3	2077	2139	0.82	0.96	0.017	0.029	0.59
CNC4	2077	2141	0.78	0.87	0.062	0.044	1.4
CNC5	2073	2137	0.57	0.91			
CNC6	2074	2136	0.68	0.97			

[a] Details regarding the parameters are presented in Table 1. Blank cells indicate parameters that have not been measured. [b] From ref *11*. [c] Values are for native sperm whale Mb and are averages from refs *10* and *12*.

Correlations between F_{in} and O_2 Binding Parameters as a Function of Distal Pocket Size. Scott et al. (*4*) analyzed the reactions of O_2 with 90 mutants of Mb and found that the bimolecular entry rate for O_2, defined as $k'_{entry} = k'_{overall}/F_{gem}$, was dependent on the volume available for ligand capture within the distal pocket and in the Xe4 cavity. To examine this idea further, we compared the measured values of F_{in} for MbCNR mutants containing small (Ala) and large (Trp or Phe) amino acids at the Leu29(B10), Val68(E11), and Ile107(G8) positions with the fractions of geminate O_2 recombination (F_{gem,O_2}), the microscopic geminate rate constants k'_{entry}, k_{bond}, and k_{escape} for O_2 binding, and the bimolecular association rate constants for NO binding, k'_{NO}, taken from Scott et al. (*4*). The oxygen binding parameters only show significant correlations with the F_{in} values for the larger isocyanides, particularly CNC5, presumably because the bound smaller isocyanides cannot sense volume changes near the back of the distal pocket (see Figures 1, 4D, and 6B).

Remarkably, there is a significant inverse correlation between F_{gem,O_2} and $F_{in,CNC5}$ (R^2 = 0.62; Figure 7A). The fraction of geminate O_2 recombination reflects the competition between internal ligand rebinding and escape to solvent on nanosecond time scales. The internal distal pocket mutations have little effect on O_2 escape (Figure 7D) but do markedly affect the speed of internal bond formation (Figure 7C). Reducing the volume sequesters the O_2 molecule close to the iron atom increasing k_{bond} markedly, whereas increasing the volume available to dissociated diatomic ligands decreases k_{bond} due an increased entropic barrier to return to the heme center (*4*). Thus, the extent of geminate O_2 recombination in I107W and V68F MbO_2 is large, whereas the F_{gem,O_2} value for L29A and I107A MbO_2 is small. The situations for L29F and V68A are more complex because of increased and decreased, respectively, direct hindrance of bound ligands.

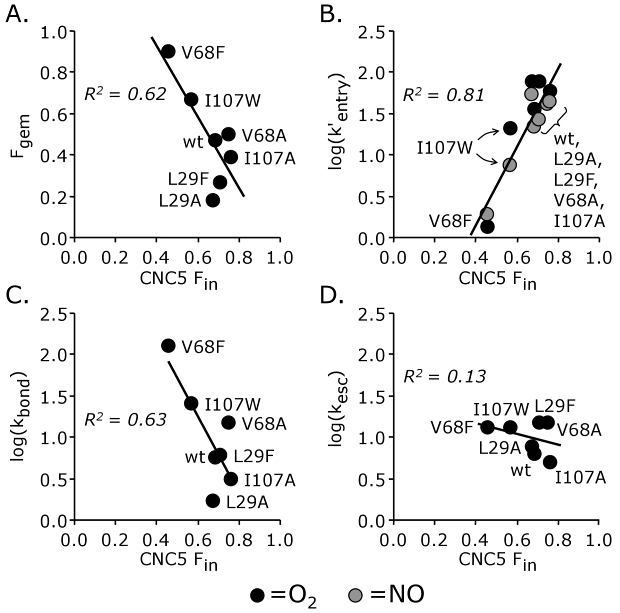


FIGURE 7: Correlations between the F_{in} values for MbCNC5 complexes and parameters describing the binding kinetics of O_2 for internal distal pocket mutants of Mb. F_{gem} and the logarithms of k'_{entry} and k_{bond}, but not k_{escape}, measured for the reaction of O_2 with Leu29, Val68, and Ile107 Mb mutants show moderate correlations with the F_{in} values for bound CNC5. k'_{NO} (gray circles, upper right panel) serves as an independent measure of the entry rate of diatomic gases. The O_2 and NO kinetic data were taken from Scott et al. (*4*).

There is a strong linear correlation between $F_{in,CNC5}$ and log(k'_{entry}) for diatomic ligands, as measured by k'_{NO} or calculated from $k'_{O_2}/F_{gem,O_2}$ (R^2 = 0.81; Figure 7B). The increase in k'_{entry} with increasing accessible distal pocket volume measured by F_{in} for MbCNC5 complexes indicates that the net rate of entry is governed by the capture volume available for the incoming ligands as well as the ease of opening of the His(E7) gate. A larger volume increases the residence time of the ligand in the pocket, allowing the His(E7) gate to close before the ligand can return to solvent, increasing the net rate of ligand capture. In contrast, there is no correlation between $F_{in,CNC5}$ and k_{escape} (Figure 7, lower right), which is governed primarily by the extent and frequency of opening of the His(E7) gate and not the volume of the distal pocket.

DISCUSSION

The results presented here, in Blouin et al. (*2*), and in Smith et al. (*1*) demonstrate that bound straight-chain alkyl isocyanides act as probes of the spaces and pathways used by ligands to bind to myoglobin. The fraction of *in* conformer obtained from FTIR measurements with wt MbCNR complexes correlates strongly and directly with the fraction of geminate recombination of bound methyl through *n*-hexyl isocyanide (CNC1−CNC6) after nanosecond laser photolysis (*2*). Photodissociated alkyl isocyanides originally oriented in the *in* conformation are trapped and efficiently rebind, but those oriented in the *out* conformation with the His(E7) gate open are much more likely to escape completely to solvent. The original evidence in support of His64(E7) regulating CNR entry and escape from Mb is shown in Figure 8. The logarithms of the measured overall bimolecular rate constants for O_2, methyl, ethyl, *n*-propyl, and *n*-butyl isocyanide binding are plotted versus the size of the amino acid at the 64(E7) position. Most of the data were taken from Rohlfs et al. (*11*) and Scott et al. (*4*), and new parameters for CNR binding to H64W Mb were measured for this work. There is an ∼10-fold decrease in k'_{O_2} when the size of the E7 amino acid is increased from Gly to Trp, but the increase in the size of the E7 amino acid has a much more dramatic effect on k'_{CNR}, which decreases from ∼100 to ≤0.1 $\mu M^{-1}\ s^{-1}$ for the same set of mutations (Figure 8).

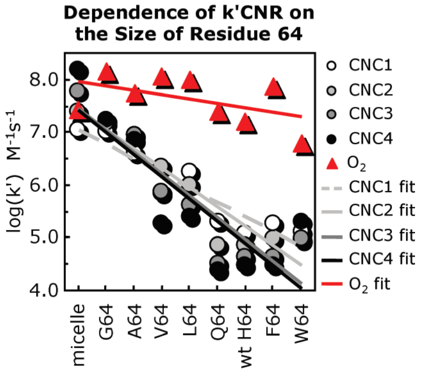
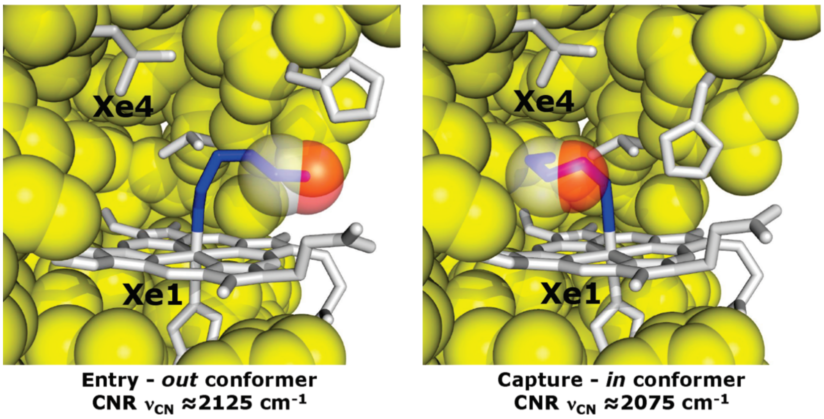


FIGURE 8: The bimolecular association rate (k') for O_2 and CNC1−CNC4 binding to Mb mutants with varied residue sizes at the E7 gate. There is only a small dependence of k'_{O_2} on the size of the residue at 64(E7). The association rate constants for the larger CNRs, however, are highly dependent on the amino acid size in the channel normally occupied by the E7 gate. The k' values are from this work (W64), Rohlfs et al. (*11*), and Olson et al. (*26*).

The complex correlations between the O_2 binding parameters, k'_{entry,O_2}, and F_{gem,O_2} and the fraction of *in* MbCNR conformer shown in Figures 3 and 7 support our idea that the alkyl side chains of isocyanide ligands act as transition state analogues for both O_2 entry and O_2 capture in the back of the distal pocket. The crystal structures of MbCNC4 taken from the second paper of this series are shown in Figure 9 and can be used to visualize diatomic ligand pathways. Two key caveats to using the isocyanides as probes of ligand channels are (1) the isocyanide alkyl chain is less mobile than a freely diffusing diatomic gas because it is tethered to the iron atom, and (2) the connectivity of the alkyl side chain may require displacement of several amino acids simultaneously rather than just one or two side chains. However, the flexibility of the straight-chained CNRs allows more accurate sensing of steric constraints and the relative free energies required for opening potential pathways than the more rigid aromatic rings of bound imidazole and phenyl ligands.

In the E7 gate mechanism, ligands enter Mb when the imidazole side chain of His64 has rotated outward, opening a direct channel to the distal pocket. This conformation is seen for the *out* conformer of native MbCNC4 $P2_1$ crystals with the *n*-butyl chain occupying an open E7 channel with enough space for a diatomic gas molecule to enter the protein (Figure 9, left panel). This interpretation is strongly supported by both the inverse and direct correlations, respectively, of $\log(k'_{entry,O_2})$ and F_{gem,O_2} with the fraction of CNC5 *in* conformers, $F_{in,CNC5}$, measured by FTIR spectroscopy for a series of mutations at the 45(CD3) and 46(CD4) positions that alter the flexibility of the His64(E7) side chain (Figure 3). In both cases, smaller values of F_{in} for *n*-pentyl isocyanide indicate a lower barrier to opening of the E7 channel, which increases the rate of bimolecular entry and at the same time decreases geminate recombination after laser photolysis by facilitating escape.

The side path or baseball glove mechanism for O_2 binding also predicts that the rate of ligand entry depends on the size of the distal pocket. The larger the volume available to incoming O_2, the



Entry - *out* conformer
CNR ν_{CN} ≈2125 cm^{-1}

Capture - *in* conformer
CNR ν_{CN} ≈2075 cm^{-1}

FIGURE 9: Left panel: Structure of the native *out* MbCNC4 conformer (PDB ID 104m (*1*)) with CO atoms placed at the terminal carbon atoms of the alkyl side chain. Right panel: Structure of the wt *in* MbCNC4 conformer (PDB ID 111m (*1*)) with CO placed in the position found in the structure of the room temperature B-state photoproduct of wt MbCO (PDB ID 1abs (*27*)). All structures were globally aligned by Cα atoms and rendered in the same orientation in PyMOL. The left panel represents ligand entry through the E7 channel, and the right panel represents ligand capture in the side path model of ligand binding first described in detail by ref *4*.

greater will be the chance of ligand capture, just as a large pocket in a baseball glove makes it easier to catch a ball. This interpretation is supported by the direct correlation between F_{in} for bound CNC5 and $\log(k'_{\mathrm{entry,O_2}})$ for mutations in which the volume of the interior of the distal pocket is increased or decreased by Ala or Phe/Trp mutations, respectively. An increase in the fraction of *in* CNC5 conformer indicates more free space in the interior of the distal pocket, and this increase in volume correlates directly with greater rates of O_2 entry (Figure 7B). This interpretation is shown in the right panel of Figure 9, where the structure of the *in* conformer of wt MbCNC4 is compared to the structure of the initial B-state structure of photodissociated CO. The terminal atom of the *n*-butyl side chain is pointing toward the Xe4 pocket which may also be part of the capture volume and is the location of state C for dissociated CO (*4*, *23*, *24*, *28*−*31*).

There is debate over the functional significance of the interior Xe cavities that were characterized at high resolution in Mb by Tilton et al. (*32*). Their original observations were followed by molecular dynamics simulations that suggested a "multiple pathways" model in which ligands can freely migrate from one Xe site to another in the interior of Mb and then exit from these positions through several transient openings to solvent, with the distal histidine gate being only a very minor pathway (*33*). Even the most recent molecular dynamics simulations (*34*−*36*) and energy mapping computations (*37*) still consistently find multiple apolar pathways for ligand entry and exit that lead from the distal pocket through the protein interior to solvent.

In 1994, Huang and Boxer (*38*) carried out a random mutagenesis study of point mutations of recombinant Mb. In their experiments, changes in O_2 and CO rebinding time courses were measured after laser photolysis of bacterial cell lysates without purifying the Mb variants. Most of the mutations that markedly affected ligand binding rates were located at or near the His(E7) position, but a number of replacements remote from the heme pocket also gave large effects. The latter observations led Huang and Boxer (*38*) to propose that their results supported the multiple pathway model.

Building on the work of Huang and Boxer (*38*), Scott et al. (*4*) used a rational mutagenesis strategy to systematically introduce small or large amino acids at positions along all of the proposed ligand entry and exit pathways in sperm whale Mb. In their work, all of the variants were expressed and purified to homogeneity, and both geminate and overall association and dissociation rate parameters were measured. They only found significant effects on rates of ligand entry and escape for substitutions near His64 and the E7 channel and suggested that ∼80% of diatomic gases enter and exit through this pathway. Some of the remote mutations reported to have large effects by Huang and Boxer (*38*) were examined and found to be false positives (*4*). Many of the other remote mutations identified in the screen of bacterial lysates involved proline substitutions in the middle of helices or other replacements that are predicted to markedly decrease Mb stability. As a result, the lysates probably contained significant amounts of partially denatured protein and free heme, which would give large changes in the time courses for CO and O_2 rebinding after laser photolysis.

Time-resolved X-ray crystallography has also provided strong experimental evidence in favor of the side path/E7 gate model proposed by Scott et al. (*3*). In crystals of wt Mb, the electron density for photolyzed CO is found to persist in the Xe1 cavity on the proximal side of the heme group with a lifetime of ∼10−20 μs and then disappears as the ligand moves out into solvent (*30*). This observation first led workers to propose that CO escapes directly to solvent from the Xe1 cavity, thus supporting the multiple paths model. However, blocking access to the distal pocket from these internal cavities by L29F and L29W mutations traps CO within the Xe1 cavity for even longer times, ∼300 and ∼1500 μs, respectively (*24*). If ligands could escape directly from the Xe sites, inhibiting their return to the distal pocket by the large Phe(B10) and Trp(B10) side chains should not have affected the lifetimes of ligand electron density in the Xe1 cavity. The simplest interpretation of these time-resolved crystallographic data is that ligands have to return to the active site to escape through the E7 channel as predicted by the side pathway model. Thus, there is still a strong disconnect between theoretical simulations that consistently predict multiple pathways and the experimental effects of mutagenesis on ligand binding that strongly identify the E7 gate and channel as the primary route for entry and exit.

In this work, we have taken another experimental approach, used CNRs to probe the spaces and paths near the Mb binding pocket, and found correlations that also indicate that ligands primarily enter and exit the protein through the His(E7) gate. An alternative explanation for the direct correlation between $F_{in,\mathrm{CNC5}}$ and $\log(k'_{\mathrm{entry,O_2}})$ is that when the connection between the Xe4 cavity and the distal pocket is blocked by the V68F and I107W mutations, ligand entry and escape through the multiple interior pathways will be markedly inhibited. However, this argument and the multiple interior pathway mechanism cannot explain the lack of correlation of $F_{in,\mathrm{CNC5}}$ with $k_{\mathrm{escape,O_2}}$ for the "pocket volume" mutants including V68F and I107W (Figure 7D). In addition, the multiple pathway model cannot readily explain the decrease in k' with increasing size of the E7 amino acid shown in Figure 8. Thus, in our view, the E7 gate/baseball glove model provides the best structural explanation of ligand binding to Mb, and bound alkyl isocyanides serve as excellent probes of the flexibility of the distal histidine gate and the capture volume in the interior of the active site, as shown in Figures 1 and 9.

REFERENCES

1. Smith, R. D., Blouin, G. C., Johnson, K. A., Phillips, G. N., Jr., and Olson, J. S. (2010) Straight-chain alkyl isocyanides open the distal histidine gate in crystal structures of myoglobin. Biochemistry (DOI: 10.1021/bi1001739).
2. Blouin, G. C., and Olson, J. S. (2010) The stretching frequencies of bound alkyl isocyanides indicate two distinct ligand orientations within the distal pocket of myoglobin. Biochemistry (DOI: 10.1021/bi100172c).
3. Scott, E. E., and Gibson, Q. H. (1997) Ligand migration in sperm whale myoglobin. *Biochemistry 36*, 11909–11917.
4. Scott, E. E., Gibson, Q. H., and Olson, J. S. (2001) Mapping the pathways for O_2 entry into and exit from myoglobin. *J. Biol. Chem. 276*, 5177–5188.
5. Olson, J. S., Soman, J., and Phillips, G. N., Jr. (2007) Ligand pathways in myoglobin: a review of Trp cavity mutations. *IUBMB Life 59*, 552–562.
6. Springer, B. A., and Sligar, S. G. (1987) High-level expression of sperm whale myoglobin in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A. 84*, 8961–8965.
7. Carver, T. E., Brantley, R. E., Jr., Singleton, E. W., Arduini, R. M., Quillin, M. L., Phillips, G. N., Jr., and Olson, J. S. (1992) A novel site-directed mutant of myoglobin with an unusually high O_2 affinity and low autooxidation rate. *J. Biol. Chem. 267*, 14443–14450.
8. Casanova, J., Jr., Schuster, R. E., and Werner, N. D. (1963) Synthesis of aliphatic isocyanides. *J. Chem. Soc.,* 4280–4281.
9. Reisberg, P. I., and Olson, J. S. (1980) Equilibrium binding of alkyl isocyanides to human hemoglobin. *J. Biol. Chem. 255*, 4144–4150.

10. Mims, M. P., Porras, A. G., Olson, J. S., Noble, R. W., and Peterson, J. A. (1983) Ligand binding to heme proteins. An evaluation of distal effects. *J. Biol. Chem. 258*, 14219–14232.
11. Rohlfs, R. J., Mathews, A. J., Carver, T. E., Olson, J. S., Springer, B. A., Egeberg, K. D., and Sligar, S. G. (1990) The effects of amino acid substitution at position E7 (residue 64) on the kinetics of ligand binding to sperm whale myoglobin. *J. Biol. Chem. 265*, 3168–3176.
12. Stetzkowski, F., Cassoly, R., and Banerjee, R. (1979) Binding of alkylisocyanides with soybean leghemoglobin. Comparisons with sperm whale myoglobin. *J. Biol. Chem. 254*, 11351–11356.
13. Carver, T. E., Olson, J. S., Smerdon, S. J., Krzywda, S., Wilkinson, A. J., Gibson, Q. H., Blackmore, R. S., Ropp, J. D., and Sligar, S. G. (1991) Contributions of residue 45(CD3) and heme-6-propionate to the biomolecular and geminate recombination reactions of myoglobin. *Biochemistry 30*, 4697–4705.
14. Lecomte, J. T., and La Mar, G. N. (1985) 1H NMR study of labile proton exchange in the heme cavity as a probe for the potential ligand entry channel in myoglobin. *Biochemistry 24*, 7388–7395.
15. Lai, H. H., Li, T., Lyons, D. S., Phillips, G. N., Jr., Olson, J. S., and Gibson, Q. H. (1995) Phe-46(CD4) orients the distal histidine for hydrogen bonding to bound ligands in sperm whale myoglobin. *Proteins 22*, 322–339.
16. Smith, R. D. (1999) Biochemistry & Cell Biology, Doctoral Thesis, p 203, Rice University, Houston, TX.
17. Egeberg, K. D., Springer, B. A., Sligar, S. G., Carver, T. E., Rohlfs, R. J., and Olson, J. S. (1990) The role of Val68(E11) in ligand binding to sperm whale myoglobin. Site-directed mutagenesis of a synthetic gene. *J. Biol. Chem. 265*, 11788–11795.
18. Quillin, M. L., Li, T., Olson, J. S., Phillips, G. N., Jr., Dou, Y., Ikeda-Saito, M., Regan, R., Carlson, M., Gibson, Q. H., and Li, H.; et al. (1995) Structural and functional effects of apolar mutations of the distal valine in myoglobin. *J. Mol. Biol. 245*, 416–436.
19. De Angelis, F., Jarzecki, A. A., Car, R., and Spiro, T. G. (2005) Quantum chemical evaluation of protein control over heme ligation: CO/O_2 discrimination in myoglobin. *J. Phys. Chem. B 109*, 3065–3070.
20. Carver, T. E., Rohlfs, R. J., Olson, J. S., Gibson, Q. H., Blackmore, R. S., Springer, B. A., and Sligar, S. G. (1990) Analysis of the kinetic barriers for ligand binding to sperm whale myoglobin using site-directed mutagenesis and laser photolysis techniques. *J. Biol. Chem. 265*, 20007–20020.
21. Smerdon, S. J., Dodson, G. G., Wilkinson, A. J., Gibson, Q. H., and Blackmore, R. S. (1991) Distal pocket polarity in ligand binding to myoglobin: structural and functional characterization of a threonine 68(E11) mutant. *Biochemistry 30*, 6252–6260.
22. Li, T., Quillin, M. L., Phillips, G. N., Jr., and Olson, J. S. (1994) Structural determinants of the stretching frequency of CO bound to myoglobin. *Biochemistry 33*, 1433–1446.
23. Aranda, R. T., Levin, E. J., Schotte, F., Anfinrud, P. A., and Phillips, G. N., Jr. (2006) Time-dependent atomic coordinates for the dissociation of carbon monoxide from myoglobin. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 62*, 776–783.
24. Schmidt, M., Nienhaus, K., Pahl, R., Krasselt, A., Anderson, S., Parak, F., Nienhaus, G. U., and Srajer, V. (2005) Ligand migration pathway and protein dynamics in myoglobin: a time-resolved crystallographic study on L29W MbCO. *Proc. Natl. Acad. Sci. U.S.A. 102*, 11704–11709.
25. Hirota, S., Li, T., Phillips, G. N., Jr., Olson, J. S., Mukai, M., and Kitagawa, T. (1996) Perturbation of the Fe-O_2 bond by nearby residues in heme pocket: observation of νFe-O2 Raman bands for oxymyoglobin mutants. *J. Am. Chem. Soc. 118*, 7845–7846.
26. Olson, J. S., McKinnie, R. E., Mims, M. P., and White, D. K. (1983) Mechanisms of ligand binding to pentacoordinate protoheme. *J. Am. Chem. Soc. 105*, 1522–1527.
27. Schlichting, I., Berendzen, J., Phillips, G. N., Jr., and Sweet, R. M. (1994) Crystal structure of photolysed carbonmonoxy-myoglobin. *Nature 371*, 808–812.
28. Brunori, M., Vallone, B., Cutruzzola, F., Travaglini-Allocatelli, C., Berendzen, J., Chu, K., Sweet, R. M., and Schlichting, I. (2000) The role of cavities in protein dynamics: crystal structure of a photolytic intermediate of a mutant myoglobin. *Proc. Natl. Acad. Sci. U.S.A. 97*, 2058–2063.
29. Ostermann, A., Waschipky, R., Parak, F. G., and Nienhaus, G. U. (2000) Ligand binding and conformational motions in myoglobin. *Nature 404*, 205–208.
30. Srajer, V., Ren, Z., Teng, T. Y., Schmidt, M., Ursby, T., Bourgeois, D., Pradervand, C., Schildkamp, W., Wulff, M., and Moffat, K. (2001) Protein conformational relaxation and ligand migration in myoglobin: a nanosecond to millisecond molecular movie from time-resolved Laue X-ray diffraction. *Biochemistry 40*, 13802–13815.
31. Bourgeois, D., Schotte, F., Brunori, M., and Vallone, B. (2007) Time-resolved methods in biophysics. 6. Time-resolved Laue crystallography as a tool to investigate photo-activated protein dynamics. *Photochem. Photobiol. Sci. 6*, 1047–1056.
32. Tilton, R. F., Jr., Kuntz, I. D., Jr., and Petsko, G. A. (1984) Cavities in proteins: structure of a metmyoglobin-xenon complex solved to 1.9 Å. *Biochemistry 23*, 2849–2857.
33. Elber, R., and Karplus, M. (1990) Enhanced sampling in molecular dynamics: use of the time dependent Hartree approximation for simulation of carbon monoxide diffusion through myoglobin. *J. Am. Chem. Soc. 112*, 9161–9175.
34. Ruscio, J. Z., Kumar, D., Shukla, M., Prisant, M. G., Murali, T. M., and Onufriev, A. V. (2008) Atomic level computational identification of ligand migration pathways between solvent and binding site in myoglobin. *Proc. Natl. Acad. Sci. U.S.A. 105*, 9204–9209.
35. Golden, S. D., and Olsen, K. W. (2008) Use of the conjugate peak refinement algorithm for identification of ligand-binding pathways in globins. *Methods Enzymol. 437*, 417–437.
36. Elber, R., and Gibson, Q. H. (2008) Toward quantitative simulations of carbon monoxide escape pathways in myoglobin. *J. Phys. Chem. B 112*, 6147–6154.
37. Cohen, J., Arkhipov, A., Braun, R., and Schulten, K. (2006) Imaging the migration pathways for O_2, CO, NO, and Xe inside myoglobin. *Biophys. J. 91*, 1844–1857.
38. Huang, X., and Boxer, S. G. (1994) Discovery of new ligand binding pathways in myoglobin by random mutagenesis. *Nat. Struct. Biol. 1*, 226–229.