Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

How the CO in myoglobin acquired its bend: lessons in interpretation of crystallographic data

Boguslaw Stec and George N. Phillips Jr*†

Department of Biochemistry and Cell Biology, W. M. Keck Center for Computational Biology, Rice University, Houston, TX 77005, USA

† Current address: Department of Biochemistry, University of Wisconsin at Madison, 433 Babcock Drive, Madison, WI 53706, USA.

Correspondence e-mail: phillips@biochem.wisc.edu

Received 6 November 2000
Accepted 24 January 2001

Contrary to the expectation of chemists, the first X-ray structures of carbon monoxide bound to myoglobin (Mb) showed a highly distorted Fe—C—O bond system. These results appeared to support the idea of a largely steric mechanism for discrimination by the protein against CO binding, a lethal act for the protein in terms of its physiological function. The most recent independently determined high-resolution structures of Mb–CO have allowed the 25 year old controversy concerning the mode of CO binding to be resolved. The CO is now seen to bind in a roughly linear fashion without substantial bending, consistent with chemical expectations and spectroscopic measurements. Access to deposited diffraction data prompted a reevaluation of the sources of the original misinterpretation. A series of careful refinements of models against the data at high (1.1 Å) and modest resolutions (1.5 Å) have been performed in anisotropic *versus* isotropic modes. The results suggest that the original artifact was a result of lower quality crystals combined with anisotropic motion and limited resolution of the diffraction data sets. This retrospective analysis should serve as a caution for all researchers using structural tools to draw far-reaching biochemical conclusions.

In the early 1970s when X-ray methods matured to the level of being universally trusted, researchers started to tackle important biochemical problems on the structural level. One of those important questions, which was posed by the very first protein solved by X-ray crystallography (Kendrew *et al.*, 1958; Kendrew, 1963), was the question about the mechanism of discrimination and specificity in the binding of dioxygen (O_2) *versus* carbon monoxide (CO) by myoglobin and hemoglobin. It was known from model studies that the preference for binding of CO over O_2 of simple porphyrin systems was ∼25 000, while in myoglobin it was only about 25. The lowering of a very high CO-binding rate lies at the very center of the ability of larger living organisms to have a functional circulatory system for oxygen delivery.

Initial structural determinations indicated that dioxygen binds in a tilted non-linear fashion (Phillips, 1980; Phillips & Schoenborn, 1981). Subsequent binding studies with CO indicated a similar bent mode of binding (Cheng & Schoenborn, 1991; Kuriyan *et al.*, 1986). This created a considerable controversy because porphyrin model compounds had never been seen to bind CO in a highly bent mode. Based on model compounds, Collman *et al.* (1976) came up with the idea (the steric hindrance hypothesis) that CO binding is much weaker because it is forced to be bent into unfavorable orientation by the protein matrix. This hypothesis was supported by what was, at the time, a high-resolution X-ray analysis. Widely accepted, the hypothesis found its place in popular biochemistry textbooks (Stryer, 1995) despite objections by spectroscopists (Li & Spiro, 1988). Nevertheless, in recent years those opposing this interpretation on chemical grounds were partially vindicated by new high-resolution (∼1.1 Å) crystal structures (Vojtechovsky *et al.*, 1999; Kachalova *et al.*, 1999).[1] The structures confirmed the results obtained from high-resolution spectroscopic methods (Lim *et al.*, 1995; Spiro & Kozlowski, 1998), indicating a mostly linear mode of CO binding. The results of Phillips *et al.* (1999), which implicate electrostatics in the discrimination process, and the recognition by spectroscopists that transition dipoles need not lie exactly along bond vectors (Spiro & Kozlowski, 1998) have allowed for a new paradigm to emerge of electrostatic stabilization by the active-site pocket with strong contribution by the distal histidine (His64) (Phillips *et al.*, 1999). We should understand, however, that strain stored in the globin could also be a partial factor in

[1] For simplicity, we will be using the abbreviated form of naming the structures in which that of Vojtechovsky *et al.* (1999) will be referred to as Berendzen's structure and that of Kachalova *et al.* (1999) as Bartunik's structure.

the discrimination process (Kachalova *et al.*, 1999; Spiro & Kozlowski, 2001; see also comments by Borman, 1999).

Motivated to learn the limitations of this generally quite powerful and objective method, we set out to reanalyze the crystallographic results obtained in the 1970s and 1980s, aiming to identify the source of the original misinterpretation. Availability of the high-resolution diffraction data for the recently published structures enabled us to study this problem in more detail. We have extracted from the Protein Data Bank the high-resolution models (1bzr and 1a6g) as well as the X-ray intensity data for both the recent analyses (Vojtechovsky *et al.*, 1999; Kachalova *et al.*, 1999). Both the former (Bartunik's) and the latter (Berendzen's) crystal structures were found to crystallize in the space group $P2_1$, the same space group reported by Kuriyan *et al.* (1986) and the original studies of Kendrew *et al.* (1958). We performed a series of refinements with *SHELXL*97 (Sheldrick & Schneider, 1997) at different resolutions and in different modes, aiming at reproducing the conditions which might initially have led to the errant conclusions reached more than 20 years ago and at learning the limits of the current data.

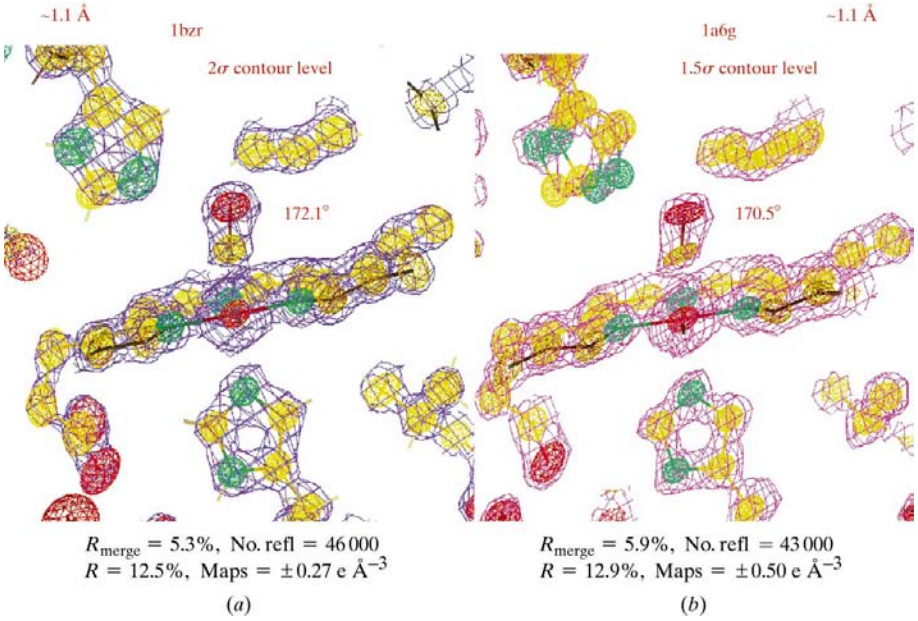

Figure 1
(*a*) The side view of the active site of sperm whale myoglobin (as determined by Bartunik and coworkers) with thermal vibration ellipsoids. Only a single CO molecule was refined. (*b*) The same view as above of the structure by Berendzen and coworkers. Note the correlation of the directions of disordered residues with the elongation of the CO ellipsoids. All figures were produced with the program *XtalView* (McRee, 1992).

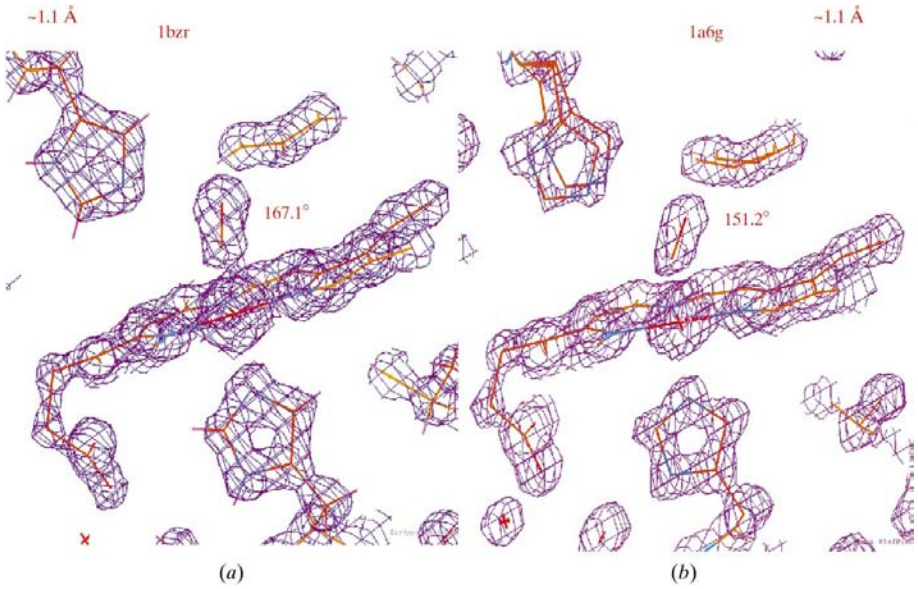

Figure 2
The side view of the active site of myoglobin refined isotropically with one CO molecule at ~1.1 Å: (*a*) the Bartunik structure, (*b*) the Berendzen structure. Note a significant bending for the CO molecule as refined against the data collected by Berendzen and coworkers. The electron-density maps are contoured at 2σ and 1.5σ, respectively.

Initially, we performed a control experiment to ascertain whether we could reproduce the results obtained by Berendzen's and Bartunik's groups. We performed ten cycles of conjugated-gradient refinements (CGLS) in *SHELXL*97 on both models in anisotropic mode with riding H atoms with the data in the 50–1.1 Å shell (all data). We were able to reproduce quite accurately the reported R values as well as the geometry of the bound CO, despite the fact that the reported protocols were substantially different from ours (Kachalova *et al.*, 1999; Vojtechovsky *et al.*, 1999). The refinement results are collected in Table 1 and Figs. 1(*a*) and 1(*b*). Refinements of both models in an anisotropic mode confirmed a nearly linear mode of binding with the C—O—Fe angle >170°. Additionally, traditional measures indicated that the quality of the Bartunik data as well as the resulting model appeared to be slightly better than that of the other group. The R_{merge} as well as the final R factors were lower in Bartunik's case. This difference was also reflected in flatter difference electron-density maps (+0.26–0.27 e Å^{-3} for 1bzr and +0.50–0.31 e Å^{-3} for 1a6g). What is more important is that this mode of refinement confirmed that Berendzen's structure had much more extensive disorder present at the active site. In particular, His64 has three distinct conformers, while in Bartunik's structure it had only one. Phe32 is also disordered, while in Bartunik's structure it appears to have a single conformation.

Subsequently, we performed a series of refinements in isotropic as well as anisotropic modes to extract detailed information about the apparent CO bending angle in the final models. After converting the deposited models to the isotropic temperature-factor representation, we performed ten cycles of conjugate-gradient refinement at ~1.1 Å resolution. The slightly higher quality of Bartunik's data was once more evident, producing a lower R factor of 18.8% as calculated on ~46 000 reflections *versus* 19.8% on ~43 000 reflections in the Berendzen structure. However, even more importantly, the Fe—C—O angle increased in both cases, to 167° in Bartunik's model

Table 1
Summary of the refinement results performed on the 1bzr and 1a6g myoglobin models against data collected to ∼1.1 Å resolution.

The data sets contained 46 002 (R_{merge} = 5.2%) and 42 857 (R_{merge} = 5.9%) reflections for 1bzr and 1a6g, respectively. The refinements were carried out on all data (50–1.1 Å) without σ cutoffs and in the anisotropic mode resulted in difference maps characterized by extreme values of +0.26–0.27 e Å^{-3} for 1bzr and +0.50–0.31 e Å^{-3} for 1a6g structures.

Model	1bzr, Bartunik	1a6g, Berendzen
Originally reported		
R factor (%)	12.4	12.8
CO bending angle (°)	171.3	170.5
10 cycles anisotropic		
R factor (%)	12.5	12.9
CO bending angle (°)	172.1	170.6
10 cycles isotropic		
R factor (%)	18.8	19.8
CO bending angle (°)	167.1	151.2
10 cycles anisotropic, 2 CO		
R factor (%)	12.6	12.9
CO bending angle (°)	171.5†	161.0, 175.8‡
10 cycles isotropic, 2 CO		
R factor (%)	—	19.8
CO bending angle (°)	—	148.1, 169.2‡
10 cycles isotropic, 1 CO, 1.5 Å resolution		
R factor (%)	15.8†	17.5†
CO bending angle (°)	165.0	140.9

† Only one CO. ‡ Two CO molecules with occupancies 0.51 and 0.49.

(Fig. 2*a*) and substantially more (to 151°) in Berendzen's model (Fig. 2*b*). Upon converting the mode of refinement back to the anisotropic mode, the Fe—C—O angle returned almost to the initial value but Berendzen's data required more cycles for the restoration of the original bend angle.

A careful inspection of the results of anisotropic refinements indicated that the ellipsoids representing CO were more elongated in the Berendzen data. The *SHELXL*97 output suggested splitting the atoms into two different sets of positions. In contrast, Bartunik's structure refined quite well isotropically without excessive elongation of thermal ellipsoids. We 'split up' the atomic coordinates for C and O atoms for further analysis of Berendzen's data, as suggested by the program results, and repeated the isotropic and anisotropic refinement with two CO molecules. The CO molecules refined to convergence, resulting in about half an ångström shift between them. Each conformer had an occupancy of approximately 50%. In the anisotropic case, the CO molecules were approximately parallel but the shift created a significant difference in orientation, with a bending angle of 176° for the first CO molecule and 161° for the second one (Fig. 3*a*). However, in the isotropic case the molecules were rotated slightly and resulted in an even larger bending angle difference, with corresponding angles of 178 and 159°, respectively.

The refinement experiments described above suggest that the degree of uniformity of the molecules in crystals leading to the deposited models was different. The Berendzen crystals apparently exhibited greater chemical inhomogeneity, which resulted in the observed conformational disorder at the active site. As a consequence, the refinement reflected those ambiguities, especially in the isotropic refinement mode. The isotropic refinement attempted to fix the disorder by producing a larger bend angle for CO. To clarify and strengthen our conclusions, we set out to imitate the conditions of the earlier determinations and decided to limit the resolution of the data to 1.5 Å. We have refined structures using both sets of diffraction data limited to 1.5 Å in the isotropic mode and with only one CO molecule present. After 20 cycles of CGLS refinement in *SHELXL*97 the bending angle of the CO molecule in Bartunik's structure remained at 165°, while in Berendzen's structure it collapsed to 141° (Fig. 3*b*). This last value corresponds very closely to the angle of 142° reported by Kuriyan *et al.* (1986) (PDB code 1mbc) for the refinement of carbonmonoxy myoglobin at 1.5 Å resolution. Therefore, the combination of the limited resolution with crystals containing much more conformational disorder can produce an artifactually bent CO conformation (Fig. 3*c*).

Incidentally, the same observations can be applied to the interpretation of the results of neutron studies (Cheng & Schoenborn, 1991). The sample was prepared under CO-binding conditions but crystals were then soaked for a very long time with D_2O to achieve full deuterium exchange with little precaution to prevent autoxidation. Mb–CO oxidizes at room temperature in a much shorter time than the time used for soaking the Mb crystals. We suggest that the sample oxidized and that the refined model represented mostly the ferric deuterium oxide Mb

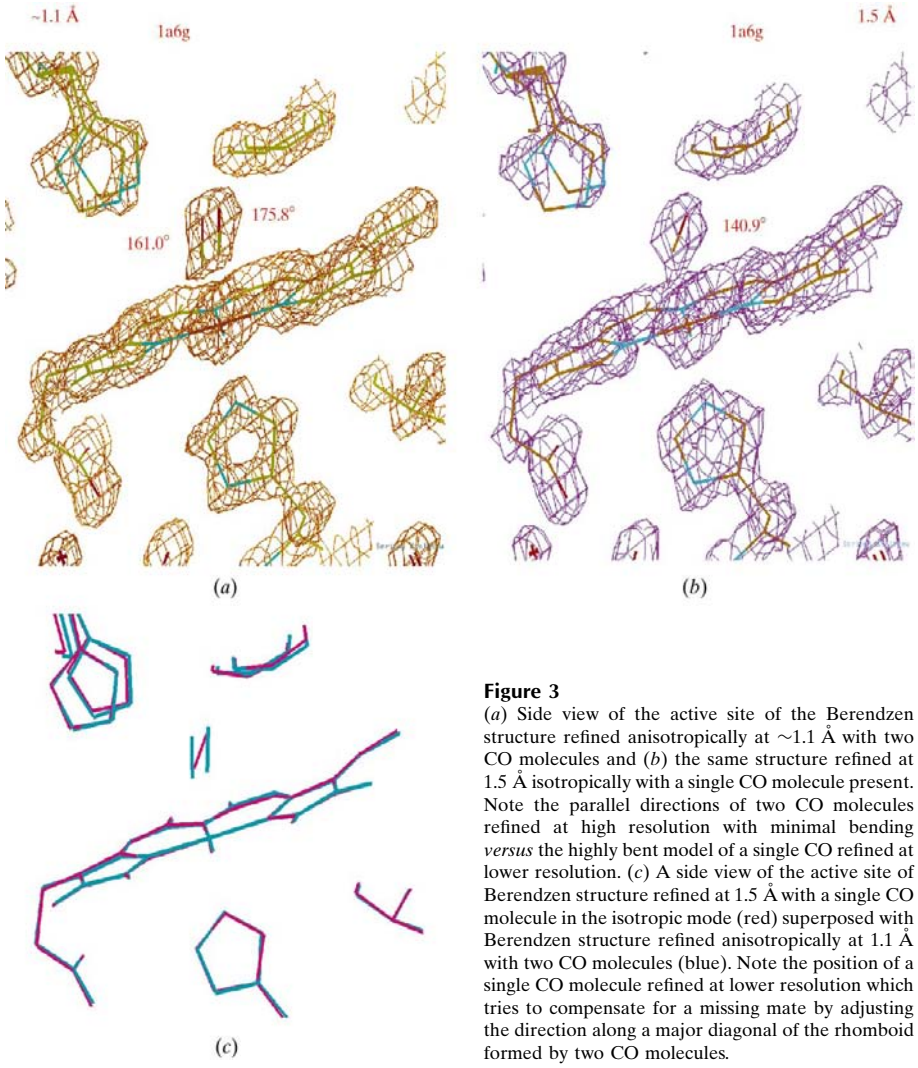

Figure 3
(*a*) Side view of the active site of the Berendzen structure refined anisotropically at ∼1.1 Å with two CO molecules and (*b*) the same structure refined at 1.5 Å isotropically with a single CO molecule present. Note the parallel directions of two CO molecules refined at high resolution with minimal bending *versus* the highly bent model of a single CO refined at lower resolution. (*c*) A side view of the active site of Berendzen structure refined at 1.5 Å with a single CO molecule in the isotropic mode (red) superposed with Berendzen structure refined anisotropically at 1.1 Å with two CO molecules (blue). Note the position of a single CO molecule refined at lower resolution which tries to compensate for a missing mate by adjusting the direction along a major diagonal of the rhomboid formed by two CO molecules.

structure. The limited resolution of such determination (∼1.8 Å) did not allow a clear distinction to be made between CO and OD^-, especially when the neutron-scattering factor for a D atom is only marginally different from that for an O atom. Additionally, the location of the D atom at the N^{δ} position reported in the original neutron studies is also in contradiction with the expected protonation state for His64 in Mb–CO (Phillips *et al.*, 1999), which provides experimental evidence in support of our interpretation.

In conclusion, we would like to suggest that (i) the quality of the crystals as measured by the amount of disorder at the active site, (ii) the effective resolution of the data and (iii) the mode of the refinement (anisotropic *versus* isotropic) has a profound influence on the refined bending angle of the CO molecule. This conclusion can be supported by the computer experiments presented in this communication, but also by comparison with the P_6 crystal form of myoglobin (Quillin *et al.*, 1993; Romo, 1998). In this crystal form, the active-site disorder seems to be less prevalent at lower resolutions of 1.6 and 1.3 Å and bending angles for CO are close to the 170° seen in the very high resolution structures.

The investigation presented should serve as a warning to all researchers using structural tools to investigate fine details of important phenomena at the atomic level. Difficult-to-control elements such as crystal quality or anisotropy of internal motions can profoundly influence the structural conclusions drawn in the presence of limited data. This is particularly important in discussing the mechanistic aspects of protein action such as catalytic mechanism and details of activation processes that depend on individual bond lengths and angles. This study also testifies to the importance of depositing the crystal diffraction data as well as models. Without the deposited data, retrospective studies such as this one would not be possible to conduct and the causes of misleading results could not be used for building experience for future studies.

We thank the National Institutes of Health (AR40252) and the W. M. Keck Center for Computational Biology (National Library of Medicine training grant LM07093) for financial support.

References

Borman, S. (1999). *Chem. Eng. News*, December, pp. 31–36.
Cheng, X. & Schoenborn, B. P. (1991). *J. Mol. Biol.* **220**, 381–399.
Collman, J. P., Brauman, J. I., Halbert, T. R. & Suslick, K. S. (1976). *Proc. Natl Acad. Sci. USA*, **73**, 3333–3337.
Kachalova, G. S., Popov, A. N. & Bartunik, H. D. (1999). *Science*, **284**, 473–476.
Kendrew, J. C. (1963). *Myoglobin and the Structure of Proteins.* Prix Nobel, 103.
Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish R. G., Wyckoff, H., Phillips, D. C. (1958). *Nature (London)*, **181**, 662–666.
Kuriyan, J., Wilz, S., Karplus, M. & Petsko, G. A. (1986). *J. Mol. Biol.* **192**, 133–154.
Li, X. Y. & Spiro, T. G. (1988). *J. Am. Chem. Soc.* **110**, 6024–6033.
Lim, M., Jackson, T. A. & Afinrud, P. A. (1995). *Science*, **269**, 962–966.
McRee, D. E. (1992). *J. Mol. Graph.* **10**, 44–47.
Phillips, G. N., Teodoro, M. L., Li, T., Smith, B. & Olson, J. (1999). *J. Phys. Chem. B*, **103**(42), 8817–8829.
Phillips, S. E. V. (1980). *J. Mol. Biol.* **142**, 531–554.
Phillips, S. E. V. & Schoenborn, B. P. (1981). *Nature (London)*, **292**, 81–82.
Quillin, M. L, Arduini, R. M., Olson, J. S. & Phillips, G. N. Jr (1993). *J. Mol. Biol.* **234**, 140–155.
Romo, T. (1998). PhD thesis, Rice University, Houston, Texas, USA.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Spiro, T. G. & Kozlowski, P. M. (1998). *J. Am. Chem. Soc.* **120**, 4524–4525.
Spiro, T. G. & Kozlowski, P. M. (2001). In the press.
Stryer, L. (1995). *Biochemistry*, 5th ed. New York: Freeman.
Vojtechovsky, J., Schlichting, I., Chu, K., Sweet, R. M. & Berendzen, J. (1999). *Biophys. J.* **77**, 2153–2175.