Structure of a de Novo Designed Protein Model of Radical Enzymes

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · OCTOBER 2002

Impact Factor: 12.11 · DOI: 10.1021/ja0264201 · Source: PubMed

J|A|C|S
COMMUNICATIONS

Published on Web 08/24/2002

Structure of a de Novo Designed Protein Model of Radical Enzymes

Qing-Hong Dai,[†] Cecilia Tommos,[‡] Ernesto J. Fuentes,[†] Margareta R. A. Blomberg,[§] P. Leslie Dutton,[†] and A. Joshua Wand[†,*]

The Johnson Research Foundation and Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6059, and Departments of Physics and Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden

Received April 4, 2002

A family of redox proteins that directly employs amino acids or posttranslationally modified amino acids as electron-transfer cofactors has grown in significance since the original proposal by Yonetani et al.[1] The use of side chains as catalytic cofactors for protein mediated redox chemistry raises significant mechanistic issues as to how these amino acids are activated toward radical chemistry in a controlled manner. Some enzymes generate and confine the radical to a single site while others allow controlled migration.[2,3] The potentially very high reactivity of the side chain free radical, combined with the apparent ability of proteins to undertake controlled migration of the radical in the protein milieu, suggests that that molecular motifs have evolved to manage this remarkable capability. It is therefore of considerable interest and importance to delineate the general structural and energetic principles by which a protein is able to promote and then control the formation of an amino acid-based radical. Mechanistic studies of side-chain redox proteins are often hampered by the properties of their radical cofactors including high reduction potentials, transient lifetimes and small optical extinction coefficients. To address some of these issues, alternative approaches have been developed that include a covalently linked ruthenium/tyrosine model complex,[4] redesign of a photosynthetic bacterial reaction center[5] and rhenium-modification of the structurally well-characterized protein azurin.[6]

We have employed *de novo* protein design to examine the structural basis for the creation and maintenance of a tryptophanyl radical.[7] A radical maquette was designed on a three-helix bundle scaffold[8] with a single tryptophan buried in the interior of the protein. To provide a stabilizing environment, the rest of the protein was designed to be redox inert. Initial characterization of the protein, termed α_3W, found it to be cooperatively stable, highly helical and to display properties consistent with the formation of a stable radical with redox potentials considerably above those seen in solution. Here we report the detailed structural analysis of α_3W using multidimensional NMR methods. The NMR work identifies a π-cation interaction that provides a structural basis for the unusual redox properties of α_3W.

The original 65-amino acid protein was recombinantly expressed in *Escherichia coli* as a fusion with thioredoxin.[9] The protein (GS-α_3W) resulting from thrombin cleavage contains a GS N-terminal extension relative to the original sequence. Studies here are based on this 67-amino acid protein, which displays all of the important characteristics of the shorter version.

* Address correspondence to this author. E-mail: wand@mail.med.upenn.edu.
† The Johnson Research Foundation & University of Pennsylvania.
‡ Department of Biochemistry & Biophysics, Stockholm University.
§ Department of Physics, Stockholm University.

Table 1. Structural Restraints and Structural Statistics

NOE-based distance restraints (1130)	
intraresidue	534
sequential	221
medium range ($1 > \|i-j\| \leq 5$)	247
long range ($\|i-j\| > 5$)	128
J-based φ angle restraints (43)	43
number of structures	30
maximum/average DYANA penalty (Å^2)	2.67/2.55
average max. distance restraint violation (Å)	0.26
average max. torsion angle restraint violation (degrees)	0.03
average pairwise rmsd (Å)	
backbone atoms (all residues)	0.76 ± 0.18
backbone atoms (residues 2−18,25−41,48−64)	0.32 ± 0.10
heavy atoms (all residues)	1.29 ± 0.11
heavy atoms (residues 2−18,25−41,48−64)	1.11 ± 0.13
Procheck-NMR[18] analysis (all residues)	
residues in most favored regions	81.4%
residues in allowed regions	16.2%
residues in generously allowed regions	1.8%
residues in disallowed regions	0.6%

The ^{15}N- and ^{13}C HSQC spectra of GS-α_3W have small chemical shift dispersions indicative of a protein with a degenerate acid content (e.g., 17 Lys, 17 Glu) and highly helical secondary structure. Nevertheless, effectively complete ^{1}H, ^{15}N, and ^{13}C resonance assignments were obtained by triple resonance and total correlation spectroscopy.[10] Stereospecific assignments of isopropyl methyl groups were obtained by the glucose-labeling approach.[11] The resonance assignments have been deposited in the BMRB databank (accession number 5356).

Though essentially complete resonance assignments could be obtained, the poor spectral dispersion made assignment of ^{1}H−^{1}H NOEs problematic. The assignment of NOEs from three-dimensional ^{15}N- and ^{13}C-resolved NOESY spectra[12] was quite limited, yielding only 550 unequivocal NOE-based distance restraints. However, four-dimensional ^{13}C,^{13}C-resolved NOESY spectra,[13] yielded an additional 580 NOE-based distance restraints. NOE-based distance restraints were supplemented with φ torsion angle restraints derived from *J*-coupling constants obtained from a 3D HNHA spectrum.[14] The experimental restraints employed are summarized in Table 1.

A family of structures was obtained using the structural refinement program DYANA.[15] The quality of the family of structural models obtained for the protein is excellent in every important respect (Table 1), indicating that a highly precise view of the protein has been obtained. The coordinates have been deposited in the PDB (accession number 1LQ7).

10.1021/ja0264201 CCC: $22.00

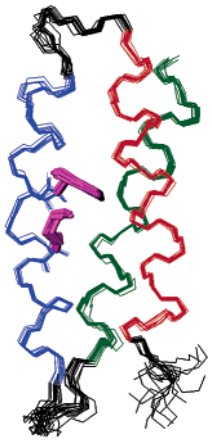

Figure 1. Main-chain superposition of the family of refined structures of GS-α_3W. Helical regions are shown in red (residues 2−18), blue (residues 25−41), and green (residues 48−64). Also shown in magenta are the heavy side-chain atoms of Trp-32 and Lys-36 that are found to form a π−cation pair.

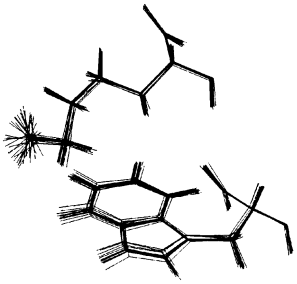

Figure 2. Side chains of Lys-36 and Trp-32 illustrating the observed π−cation interaction. Shown are local superpositions of the subset of 16 structures having Lys-36 χ_4 values centered on −89° (see text).

As designed, the protein (see Figure 1) folds into a compact three helix bundle with average inter-helical angles of ∼26° (helices 1/2), ∼21° (helices 2/3), and ∼19° (helices 1/3) and a clockwise topology leading to favorable inter-helical E−K salt bridge interactions. The indole ring of the sole tryptophan residue is largely buried within the interior of the bundle, although the indole NH does have significant solvent exposure.

The indole ring displays an orientation with the side chain of Lys-36 that is suggestive of a π−cation interaction with the ϵ-amino group. While the indole ring occupies the same orientation with respect to Lys-36 in all structures, the end of the Lys-36 side chain is distributed roughly equally (16/14) between two conformers arising from χ_4 values of $-89 \pm 2°$ and $+92 \pm 4°$. The former conformer results in a distance of 4.7 Å from the ϵ-amino nitrogen to the center of the benzoid ring of the indole side chain and an angle of 38° to the ring plane normal. These values are in the center of the distribution seen for π−cation interactions in natural proteins.[16] The other χ_4 conformer of Lys-36 that is seen in the family of structures gives parameters that are not within the ranges seen in natural proteins.

This ambiguity was resolved on the basis of the chemical shifts of the Lys-36 side chain. Although the chemical shift was not employed as a refinement parameter, an analysis of the ring current shifts[17] indicates that the conformer most consistent with the chemical shifts observed for the Lys-36 side chain is also consistent with a π−cation interaction. The orientation of the two side chains in this conformer is shown in Figure 2.

The effects of the π−cation interaction are potentially complicated. Previous work indicates that the reduced anion of the indole does not participate in the redox chemistry,[7] leaving only the neutral reduced species (WH) to be considered. Simple electrostatic arguments indicate that the cation W^oH^+/WH potential will be raised by the π−cation interaction. How a π−cation interaction will alter the potential of the neutral W^o/WH pair is less obvious and was examined by calculating the indole N−H bond strength in the presence or absence of a π−cation using an indole-methylammonium model system.[19] The N−H bond strength was found to increase by 6.8, 3.2, and 1.4 kcal/mol in the presence of the π−cation in a dielectric medium of 1, 4, and 80, respectively. Thus, these calculations are consistent with the π−cation interaction observed in the structure contributing to the measured raised reduction potential of the tryptophan.[7] In conclusion, the radical protein maquette described here provides a unique platform with which to explore the structural basis for the extraordinary radical chemistry of a critically important class of enzymes.

Acknowledgment. We are grateful to Dr. R. Koder for discussion. Supported by NIH Grants GM35940 and GM48130 and by NSF Grant DMR00-79909.

References

(1) Yonetani, T.; Schleyer, H.; Ehrenberg, A. *J. Biol. Chem.* **1966**, *241*, 3240. Reichard, P.; Ehrenberg, A. *Science* **1983**, *221*, 514.
(2) Stubbe, J.; van der Donk, W. A. *Chem. Rev.* **1998**, *98*, 705. Okeley, N. M.; van der Donk, W. A. *Chem. Biol.* **2000**, *7*, R159. Schwartz, B.; Klinman, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14766. Pesavento, R. P.; van der Donk, W. A. *Adv. Protein Chem.* **2001**, *58*, 317.
(3) For example: Tommos, C.; Babcock, G. T. *Acc. Chem. Res.* **1998**, *31*, 18. Proshlyakov, D. A.; Pressler, M. A.; DeMaso, C.; Leykam, J. F.; DeWitt, D. L.; Babcock, G. T. *Science* **2000**, *290*, 1588. Ekberg, M.; Pötsch, S.; Sandin, E.; Thunnissen, M.; Nordlund, P.; Sahlin, M.; Sjöberg, B.-M. *J. Biol. Chem.* **1998**, *273*, 21003. Becker, A.; Fritz-Wolf, K.; Kabsch, W.; Knappe, J.; Schultz, S.; Wagner, A. F. V. *Nat. Struct. Biol.* **1999**, *6*, 969. Aubert, C.; Vos, M. H.; Mathis, P.; Eker, A. P. M.; Brettel, K. *Nature* **2000**, *405*, 586.
(4) Sjödin, M.; Styring, S.; Åkermark, B.; Sun, L.; Hammarström, L. *J. Am. Chem. Soc.* **2000**, *122*, 3932.
(5) Kálmán, L.; LoBrutto, R.; Allen, J. P.; Williams, J. C. *Nature* **1999**, *402*, 696.
(6) Di Bilio, A. J.; Crane, B. R.; Wehbi, W. A.; Kiser, C. N.; Abu-Omar, M. M.; Carlos, R. M.; Richards, J. H.; Winkler, J. R.; Gray, H. B. *J. Am. Chem. Soc.* **2001**, *123*, 3181.
(7) Tommos, C.; Skalicky, J. J.; Pilloud, D. L.; Wand, A. J.; Dutton, P. L. *Biochemistry* **1999**, *38*, 9495.
(8) Johansson, J. S.; Gibney, B. R.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1998**, *120*, 3881.
(9) The amino acid sequence of GS-α_3W is: GSRVKALEEKVKALEEKVK-ALGGGGRIEELKKKWEELKKKIEELGGGGEVKKVEEEVKKLEEE-IKKL. The structure was determined using data collected at 25° C and pH 5.5. The sequence numbering of the original sequence is used throughout, i.e., the N-terminal residues GS are numbered here as −2 and −1, respectively.
(10) Wittekind, M.; Mueller, L. *J. Magn. Reson., Ser. B* **1993**, *101*, 201. Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1992**, *114*, 6291. Montelione, G. T.; Lyons, B. A.; Emerson, S. D.; Tashiro, M. *J. Am. Chem. Soc.* **1992**, *114*, 10974. Bax, A.; Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1990**, *88*, 425. Grzesiek, S.; Bax, A. *J. Magn. Reson.* **1992**, *96*, 432. Jahnke, W.; Kessler, H. *J. Biomol. NMR* **1994**, *4*, 735.
(11) Neri, D.; Szyperski, T.; Otting, G.; Senn, H.; Wüthrich, K. *Biochemistry* **1989**, *28*, 7510.
(12) Kay, L. E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663. Zhang, O.; Kay, L. E.; Olivier, J. P.; Forman-Kay, J. D. *J. Biomol. NMR* **1994**, *4*, 845. Ikura, M.; Kay, L. E.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1990**, *86*, 204. Muhandiram, D. R.; Farrow, N. A.; Xu, G. Y.; Smallcombe, S. H.; Kay, L. E. *J. Magn. Reson., Ser. B* **1993**, *102*, 317.
(13) Clore, G. M.; Kay, L. E.; Bax, A.; Gronenborn, A. M. *Biochemistry* **1991**, *30*, 12.
(14) Vuister, G. W.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 7772.
(15) Guntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, *273*, 283.
(16) Minoux, H.; Chipot, C. *J. Am. Chem. Soc.* **1999**, *121*, 10366.
(17) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graph.* **1996**, *14*, 51.
(18) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
(19) B3LYP hybrid density functional calculations were done with the program Jaguar (ver. 4.0, Schroedinger, Inc., Portland, OR). The model system was optimized in the gas phase using the 6-31G basis set. The relative energies of the optimized structures were calculated at the B3LYP level using the 6-311G** basis set. Dielectric effects were calculated using the self-consistent reaction field method of the Jaguar program.

JA0264201