

Structure, dynamics, and redox reactivity of an all-purpose flavodoxin

Received for publication, December 13, 2023, and in revised form, February 15, 2024. Published, Papers in Press, February 28, 2024.
<https://doi.org/10.1016/j.jbc.2024.107122>

Sharique Khan¹, Ahmadullah Ansari², Monica Brachi³, Debarati Das¹, Wassim El Housseini³, Shelley Minter^{3,4}, and Anne-Frances Miller^{1,*}

From the ¹Department of Chemistry, University of Kentucky, Lexington, Kentucky, USA; ²Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA; ³Department of Chemistry, University of Utah, Salt Lake City, Utah, USA; ⁴Department of Chemistry, Kummer Institute Center for Resource Sustainability, Missouri University of Science and Technology, Rolla, Missouri, USA

Reviewed by members of the JBC Editorial Board. Edited by Joan B. Broderick

The flavodoxin of *Rhodopseudomonas palustris* CGA009 (*Rp9Fld*) supplies highly reducing equivalents to crucial enzymes such as hydrogenase, especially when the organism is iron-restricted. By acquiring those electrons from photodriven electron flow *via* the bifurcating electron transfer flavoprotein, *Rp9Fld* provides solar power to vital metabolic processes. To understand *Rp9Fld*'s ability to work with diverse partners, we solved its crystal structure. We observed the canonical flavodoxin (Fld) fold and features common to other long-chain Flds but not all the surface loops thought to recognize partner proteins. Moreover, some of the loops display alternative structures and dynamics. To advance studies of protein–protein associations and conformational consequences, we assigned the ¹⁹F NMR signals of all five tyrosines (Tyr). Our electrochemical measurements show that incorporation of 3-¹⁹F-Tyr in place of Tyr has only a modest effect on *Rp9Fld*'s redox properties even though Tyrs flank the flavin on both sides. Meanwhile, the ¹⁹F probes demonstrate the expected paramagnetic effect, with signals from nearby Tyrs becoming broadened beyond detection when the flavin semiquinone is formed. However, the temperature dependencies of chemical shifts and linewidths reveal dynamics affecting loops close to the flavin and regions that bind to partners in a variety of systems. These coincide with patterns of amino acid type conservation but not retention of specific residues, arguing against detailed specificity with respect to partners. We propose that the loops surrounding the flavin adopt altered conformations upon binding to partners and may even participate actively in electron transfer.

From shortly after their initial description (1, 2) flavins were recognized as remarkable on account of their ability to execute one-electron (1e) transfers, in contrast to the 2e redox activity

of most other dyes known at the time (3). Indeed, the FMN-containing protein flavodoxin (Fld) was so named based on its ability to serve many of the roles played by ferredoxin (Fd), a [Fe₄-S₄] cluster-containing carrier of 1e equivalents at low reduction midpoint potential (E°) (4, 5). Use of an organic cofactor instead of Fe would have enabled early life to exploit environments previously uninhabitable due to Fe scarcity. Moreover, diminished reliance on Fe would have become critical when more oxidizing conditions prevailed and made Fe less bioavailable and toxic, by virtue of the Fenton chemistry Fe catalyzes. Modern organisms continue to use Fld instead of Fd under conditions of oxidative stress or Fe scarcity (6–11). Indeed, Fld is prevalent in our pathogens, enabling them to survive the hoarding used by hosts to deprive pathogens of Fe (10, 12). Conversely, augmentation of plant genomes with Fld genes has shown promise for increasing plants' tolerance to stress (13).

Flavodoxins comprise a diverse superfamily of 145 to 180 amino acid long proteins which use a noncovalently bound FMN to mediate electron transfer. (The specialized Fld-like ribonucleotide reductase–activating Nrdl (6, 14) and the more distantly related WrbA classes display diverging reactivities and are not discussed here (15).) Collectively, bacterial Flds service a wide variety of partner proteins including Fd (Fld) NADP⁺ oxido-reductase (FNR, (16, 17)), nitrogenase (18–20), nitrite reductase (21), biotin synthase (22, 23), methionine synthase (24, 25), hydrogenase (26, 27), pyruvate formate lyase (28), pyruvate oxidoreductase and 2-oxoglutarate oxidoreductase (29), cytochrome P450 (30), radical SAM enzymes (31), and likely more. Thus, individual organisms have been found with as many as 14 different Flds (10).

Based on the distributions of Fld and Fd in the tree of life, Fld is inferred to have evolved later than Fd and gained prominence following the rise in O₂ (10). However the smaller diversity of Flds *versus* Fds suggests that individual Flds may serve more partners. We infer that Fld binding should be less selective than binding of Fd, although Fld's binding motif must be similar to that of Fd, since Fld must bind the same partners. Flds are indeed found to conserve a negatively charged patch surrounding the site where the xylene ring of the FMN

* For correspondence: Anne-Frances Miller, afmill3r2@gmail.com.

Present addresses for: Ahmadullah Ansari, Center for Advanced Medicine and Biotechnology, Rutgers University, Piscataway, NJ 08854-8021; Shelley Minter, Kummer Institute Center for Resource Sustainability, Department of Chemistry, Missouri University of Science and Technology, Rolla, MO 65409.

protrudes into solution (Fig. S1, B–D) (24, 32). This resembles the negatively charged binding surface of Fd (33). However, Flds, especially long-chain Flds, are distinguished by the presence of several loops that surround the protruding flavin and have been proposed to guide Fld's association with partner proteins (12, 34).

While the first protein named 'flavodoxin' was isolated from *Clostridium* (4) and is now known to be a short chain (SC) Fld, a prior report of a protein named 'phytoflavin' can now be attributed to a long chain (LC) Fld (35). The SC and LC Flds diverged early and are distinguished by the absence or presence of a ~20 residue 'long loop' (LL) excursion from β strand 5 that buttresses the tyrosine 90 (Y90) which in turn stacks against the *si* face of the flavin ((36) numbering of *Rhodospseudomonas palustris* will be the default throughout). Thus the LL provides a flexible cowl (8, 36) that expands the protein's interactive surface surrounding the flavin.

The LL does not appear to confer stability, although it strengthens affinity for FMN (36) and may protect the flavin from oxidative stress (10). Due to its association with photosynthetic electron transfer to NADH (10, 18), its significant sequence conservation, and its interactions with partner proteins (24, 37), it is widely believed to mediate interactions with partner proteins (34). Thus, it could play a role in selectivity, location of the correct binding surface, and securing Fld in a productive pose for electron transfer. Nevertheless, in organisms which possess only a single Fld, the Fld must interact with several different partners, suggesting that adaptive loop(s) would be a better binding strategy than a fixed surface. This is the situation in the phototroph *R. palustris* CGA009 (38), so we have overexpressed its sole Fld, *Rp9Fld*, to test the hypotheses that the LL will be flexible; other loops surrounding the protruding flavin may also be flexible, but non-loop regions distant from the flavin should not be.

To characterize local dynamics, we are developing ^{19}F NMR probes of individual side chains, especially residues in the surface loops of Fld that may respond differently to different partners (24). ^{15}N NMR relaxation has shown that surface loops have lower backbone order parameters than the protein as a whole (39) but ^{15}N relaxation focuses on rapid motions in the ps - ns time scale, whereas we want to investigate the LL's role in slower processes of binding to and detaching from large partner protein complexes. For this, it is ideal to place ^{19}F in amino acid side chains believed to mediate the interactions with partner or residues that trigger binding/release. Indeed, NMR revealed two environments affecting the conserved Y90 flanking the flavin (40). However, those data also raised concerns regarding the integrity of the protein. To provide better ^{19}F NMR coverage of the Fld while retaining the critical conserved Tyr's and to assess the generality of our findings, we have now applied the same methods to Fld from a different strain of *R. palustris*. Using a combination of X-ray crystallography and NMR spectroscopy, we confirm that the loops surrounding the flavin access diverse conformations, setting the stage for adaptive binding. Moreover, a role in partner binding is supported by the analysis of >500 different Fld sequences.

Results

Crystal structure of *Rp9Fld* preserves well-known Fld fold

To exploit availability of a fully sequenced genome, the Fld from strain CGA009 of *R. palustris* (38) was employed (*Rp9Fld* hereafter, = Uniprot ID Q6N7Y7). The Fld we studied previously was from strain 2.1.18 (*Rp8Fld*) and was also the sole Fld in its genome, suggesting that binding surfaces of these two Flds should be conserved. However, these two orthologous Flds differ in 25% of their amino acids. Sixteen distinguishing residues are similar, but 24 are not (Fig. S1A).

To assess the potential significance of the differences and make sense of previous findings regarding multiple environments for conserved Y90 (40), we solved the crystal structure of *Rp9Fld*. This revealed most of the differences between the two Flds to be surface-exposed and predominantly on faces not proposed to interact with partner proteins (Fig. S1A, (41)).

We determined the structure of *Rp9Fld* at 2.1 Å resolution based on diffraction data collected at 100 K (data collection and refinement statistics appear in Table 1). *Rp9Fld* displays the signature flavodoxin fold, including the 5-stranded β sheet sandwiched between five α helices (Fig. 1); however, it has a shorter loop from 59 to 60 than the corresponding loop from 64 to 74 in *Rhodobacter capsulatus* (2WC1) or *Azotobacter vinelandii* (5K9B) (18, 42).

We modeled our Fld in terms of two protomers per asymmetric unit; chains A and B, because some loops of chain B

Table 1
X-ray diffraction data collection (in house) and refinement statistics

Wavelength (nm)	1.5406
Resolution range (Å) ^a	24–2.17 (2.248–2.17)
Space group	P 1 21 1
Unit cell (Å) a b c	46.6639 58.8915 60.9043
Unit cell (°) α β γ	90105.494 90
Total reflections ^a	131,196 (12,875)
Unique reflections ^a	16,966 (1672)
Multiplicity	7.7 (7.7)
Completeness (%) ^a	99.80 (99.76)
Mean I/sigma(I) ^a	22.72 (4.20)
Wilson B-factor	30.85
R-merge ^a	0.09818 (0.5687)
R-meas ^a	0.1055 (0.6112)
R-pim ^a	0.03776 (0.2208)
CC1/2 ^a	0.983 (0.932)
CC ^a	0.996 (0.982)
Reflections used in refinement ^a	16,947 (1668)
Reflections used for R-free ^a	825 (78)
R-work ^a	0.2133 (0.2612)
R-free ^a	0.2689 (0.3554)
CC(work) ^a	0.944 (0.876)
CC(free) ^a	0.907 (0.748)
Number of non-H atoms	2605
Macromolecules	2439
Ligands	62
Solvent	104
Protein residues	329
RMS(bonds)	0.008
RMS(angles)	0.91
Ramachandran favored (%)	95.22
Ramachandran allowed (%)	4.78
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.40
Clashscore	8.57
Average B-factor	35.90
Macromolecules	35.90
Ligands	37.63
Solvent	34.95

^a Statistics for the highest-resolution shell are shown in parentheses.

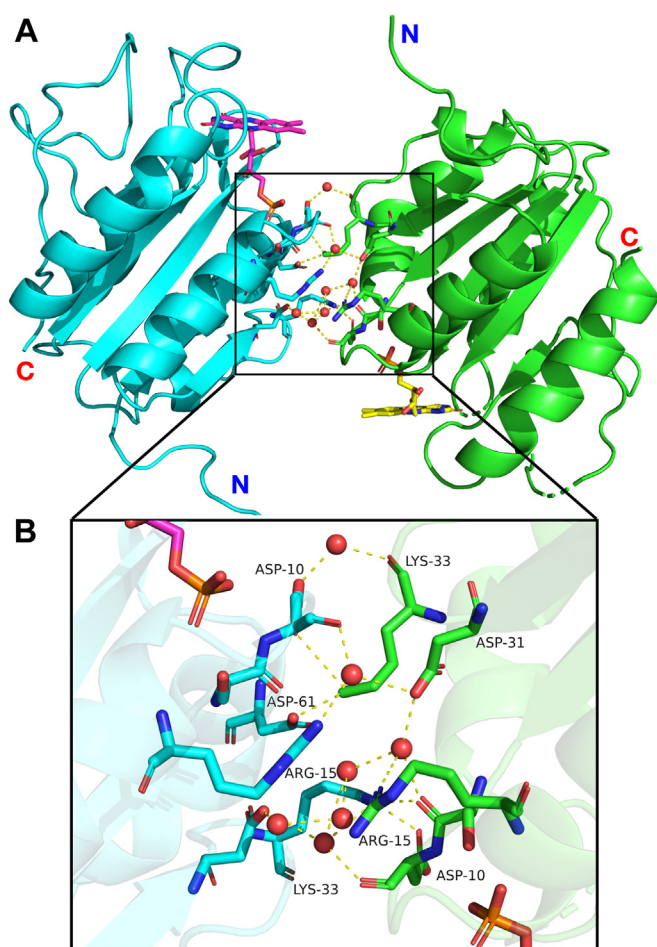


Figure 1. Ribbon structure of *Rp9Fld*, showing the opposing orientations of the two protomers and the interface between them. Panel (A) shows complete ribbon diagrams of both protomers, with unresolved backbone of protomer (B) indicated by dashed lines. The long N termini include a resolved AARGSH peptide representing the LVPRGSH of the thrombin cleavage motif that immediately precedes the natural N terminus of *Rp9Fld* in our construct. Panel (B) shows the interactions that bridge the protomers, in this relatively polar interface. The B chain is shown on the right. Figure was generated using Pymol (<http://www.pymol.org>).

have C α s displaced by as much as 4.0 Å from their positions in chain A (RMSD = 0.739 for all C α) (Fig. 1). The two protomers are oriented in opposite directions (the flavin appearing near the top for one protomer when the flavin of the other appears near the bottom, Fig. 1A). The 726 Å² interface between protomers constitutes some 10% of the surface area of one protomer based on PDBEPIA (43) and is stabilized by 12 hydrogen bonds involving primarily residues from helix 3. Helix 3-mediated association between oppositely oriented protomers is also observed between the two Fld domains in dimeric flavodiiron proteins, where dimerization brings together the FMN of one protomer with the diiron site of the other, enabling electron transfer between them (Fig. S2, Moorella, 1YCH and 4D02, (44, 45)). However, consistent with the relatively small interface area shared by its two protomers, *Rp9Fld* is a monomer in solution, based on size-exclusion chromatographic elution as a single species with an effective mass of 17.5 kDa $\pm$ 5%.

Signatures of disorder in crystals at two temperatures

A second distinction between the two protomers in *Rp9Fld*'s crystallographic unit cell is that stretches of chain B could not be resolved whereas the same residues in chain A were well ordered (Fig. 2A). This indicates that position in the crystal lattice plays a role. Disorder in the loop containing S146 of chain B may reflect a clash between S146.B and E63.B in an adjacent unit cell. Other unresolved regions in chain B adjoin cavities in the reconstructed lattice, so we speculate that these stretches populate multiple local conformations. Both phenomena could reflect loops that are dynamic and/or adopt a variety of conformations as part of crystallization. To distinguish between mobility *versus* static disorder as a basis for the

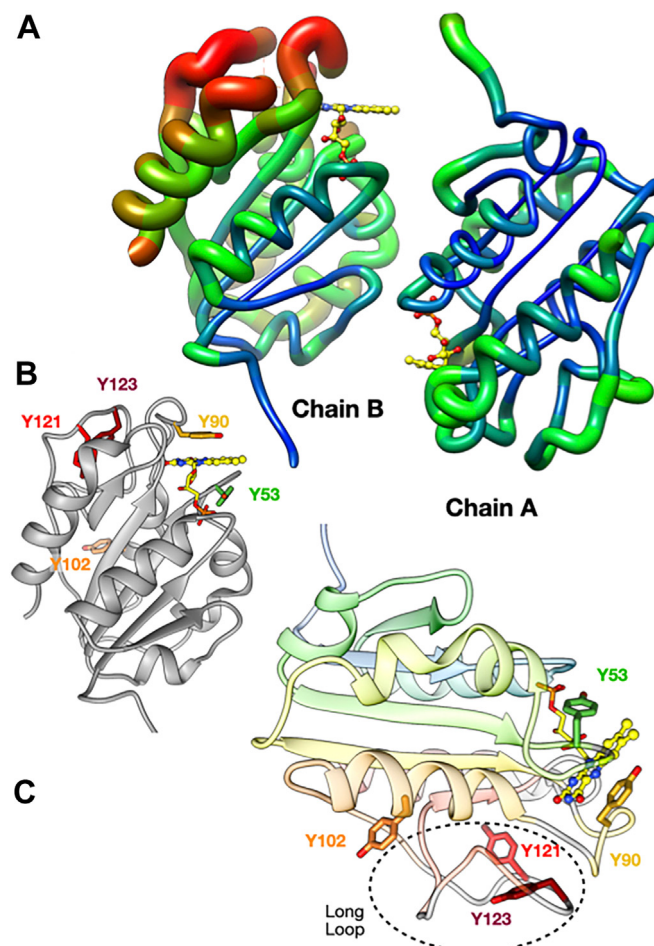


Figure 2. Structure and disordered regions of *Rp9Fld*. A, B-factors are mapped onto a cartoon of the Fld structure with wide worms (2 Å) representing regions with the highest B-factors (71 Å²) and narrow worms (0.25 Å) representing regions with the lowest B-factors (19 Å²). A red-green-blue scale reinforces the same information with highest B-factors in red. Panel (B) reproduces the same orientation used for the B chain in panel (A) but using the A chain of the low-temperature structure instead to provide the complete backbone and displaying the Tyr side chains colour-coded as in spectra that follow. C, the ribbon topology of the A chain is shown via a blue-to-red N-to-C rainbow in which gray is used to indicate residues that are not resolved in the B chain at RT, residues 54, 86, 93 to 94, 117 to 125, 130 to 132, 141 to 148. The LL is indicated with a dashed oval and contains Y121 and Y123. The loop containing Y90 is sometimes called the Y-loop, while the loop containing Y53 in our case, but W in many others, is known as the W-loop. These molecular graphics and those in later figures were made using Chimera (88, 100). LL, long loop.

low resolution of these loops, we used another crystal to collect diffraction data at RT, with the understanding that at this higher temperature, mobility would be greater due to higher thermal energy.

The structure based on RT data was lower in resolution overall (2.9 Å) and included more unrefinable stretches than the structure based on data collected at 100 K (Fig. 2, A and C). All of these were still in chain B though, and the fully resolved A chain of the RT structure overlayed well with the chain A based on 100 K data (RMSD = 0.447 Å across 165 pairs for chains A *versus* 1.718 Å across 134 pairs for chains B). Thus, we interpret the RT model as an additional perspective on the same structure, but one in which local structure is more mobile and/or more amenable to distortion by crystal packing forces. In Figure 2C, regions of chain B that are unresolved at RT are in gray. Most of the chain B residues that were disordered at 100 K were also disordered in the RT data along with numerous additional residues (exceptions are 56, 87–92, and 129). This temperature dependence suggests that causes of the high B-factors and unrefined stretches include dynamics, not just static disorder. The sections not resolved at RT are interesting because three of them occurred in loops observed to dock with the photosystem I protein in the cryo-EM structure of a Fld•PSI complex (37) and/or to have altered NMR chemical shifts upon binding to methionine synthetase and FNR partner domains (24). Thus, the dynamics we observe in our crystal structures could have functional significance.

Validation of ^{19}F NMR probes of environment and dynamics

We used NMR spectroscopy to test for conformational dynamics in solution. To compare regions of the protein that

were different in the two monomers or at different temperatures, we targeted Tyr side chains. Incorporation of 3- ^{19}F -Tyr enables observation of Tyr side chains *via* highly sensitive ^{19}F NMR, with essentially no spectral background. We accomplished this by producing R_p9Fld in bacterial growth medium containing 3- ^{19}F Tyr instead of Tyr, while inhibiting *de novo* Tyr biosynthesis. R_p9Fld has a useful number of Tyr side chains distributed in the regions of interest. Y121 and Y123 reside in a region that was unresolved by crystallographic data obtained at RT; Y53 is immediately adjacent to the flavin on its *re* face and Y90 is stacked against the flavin on its *si* face (Fig. 2C). Y102 provides a valuable control, reporting on integrity of the overall structure, as it is in the center of a helix that is well resolved in the crystal structures obtained at both temperatures and not close enough to be directly influenced by the flavin.

The ^{19}F NMR spectrum is fully resolved but displays *six* strong resonances, although there are only *five* Tyrs (Fig. 3, bottom spectrum). However, two of the resonances have integrals only half as large as those of the other signals (Table 2). In the R_p8Fld, a pair of half-strength resonances was also observed and was assigned to Y90 (40), so we determined whether the same is true in this case.

We assigned all the resonances by generating a series of singly mutated *fldA* genes, to remove each Tyr one at a time *via* single site substitutions (Fig. 3). For each position, we converted the existing Tyr to a residue that commonly replaces it among other LC Flds (Table S1). All the variant Flds were well-behaved and one resonance disappeared altogether while others barely shifted, permitting assignment of the absent resonance to the Tyr that had been changed (Fig. 3). Both the

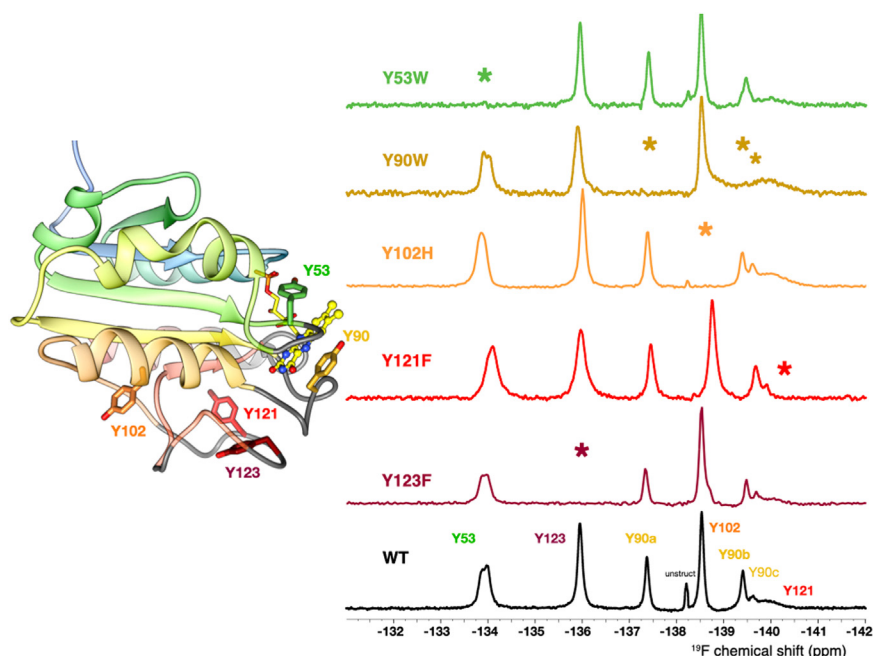


Figure 3. NMR spectrum assigned *via* replacements of individual Tyrs. In each spectrum, the location of a missing resonance is indicated with an asterisk. The WT resonance at this position is thereby assigned to the residue that had been replaced in the variant. In the case of Y121, replacement with F also caused a slight shift in the resonance of Y123. Spectra were processed with 50 ms Gaussian weighting to maximize signal-to-noise because the samples of amino acid-substituted variants were more dilute (weaker signals) than those of WT R_p9Fld. A ribbon structure is provided for reference, colored with a blue-yellow-red rainbow (N- to C-terminus) to depict primary structure.

Table 2
Chemical shifts, integrals, and line widths *versus* temperature

Residue	Chemical shift @ 25 °C (ppm ^a)	Integral area <i>versus</i> Y123 @ 6 °C	Deconvoluted area <i>versus</i> Y123 @ 6 °C	Temperature dependence of chemical shift rel. to unstructured Y. (ppb/°C)
Y53	-133.982	1.039	1.09	0
Y90a	-137.381	0.525	0.53	-4
Y90b	-139.424	0.379	0.35	7
Y90c	-139.64		0.03	9
Y102	-138.545	1.038	1.05	2
Y121	-139.9	0.691	0.80	29
Y123	-135.967	1.00	1.0	2
Unstructured Y	-138.223	0.183	0.16	-

^a Referenced to KF at -121 ppm.

half-strength resonances were absent from the Y90W variant, along with a smaller sharp resonance near the upfield half-strength resonance (resonances Y90 a, b, c in Table 2). All three were therefore assigned to Y90 and the fact that there are several indicates that FldA populates two dominant conformations that affect Y90, plus a third minority conformation, at least.

Because it is unusual to lose multiple signals at once upon substitution of a single amino acid, we applied a second test of these signal's assignment to Y90. Y90 and Y53 are immediately adjacent to the flavin ring based on our crystal structure, so both should be paramagnetically relaxed when Fld adopts its semiquinone state. In contrast, all other Tyr side chains are ≥ 8.5 Å away. Indeed, Figure 4 shows that formation of the neutral semiquinone (NSQ) state made signals attributed to both Y53 and Y90 invisible, presumably due to extreme line broadening. The NSQ was formed by titration with dithionite, as confirmed based on the sample turning blue. The fact that none of the other resonances shifted significantly indicates that flavin reduction did not change the structure in a way affecting environments of Tyr 102, 121, or 123. The linewidths of Tyr 102, 121, or 123 were also unchanged. The fact that Y121's resonance retained the same breadth, and indeed remained visible at all, indicates that its breadth reflects dynamics rather than proximity to paramagnetic flavin. Therefore, based on their insensitivities to flavin oxidation state, we can attribute

the different linewidths of Y102 and Y123 *versus* Y121 to different dynamics instead.

The assigned resonance positions reveal interesting features. Y53 displays a significant downfield shift relative to the other resonances (Fig. 3), consistent with hydrogen bond donation from the phenolic alcohol to the FMN phosphate. Additionally, the resonances vary considerably in width at half-height (Fig. 3), suggesting relaxation due to internal dynamics in addition to overall tumbling of the Fld. Thus, our ¹⁹F spectrum reports on both residue-specific environments and dynamics that vary over the protein.

Distinguishing slow- versus fast-exchange dynamics via dependence on temperature

Multiple co-existing conformations can produce sharp NMR lines if the conformations interconvert so slowly that they behave like non-interconverting structures (slow-exchange limit) or if interconversion is so rapid that NMR perceives a single average conformation and reports a single chemical shift that is the population-weighted average of those of the participating states (fast exchange limit). To test whether the multiple sharp lines attributed to Y90 represent alternative conformations that interconvert slowly on an NMR timescale and whether any other lines might reflect multiple conformations interconverting, we modulated the rate of interconversion by lowering and raising the temperature.

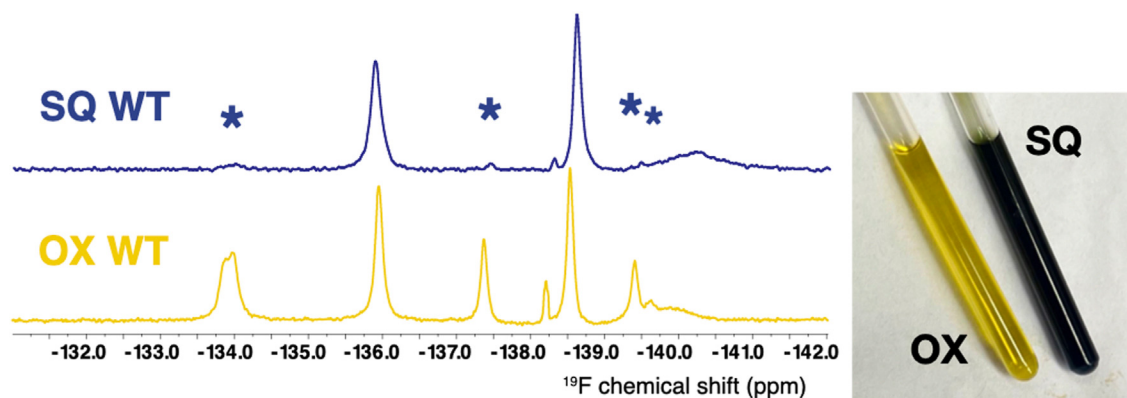


Figure 4. Extreme paramagnetic broadening due to FMN semiquinone identifies resonances of residues nearby. The NMR spectrum of oxidized (OX) *Rp9Fld* in black is compared with that of the NSQ state (red spectrum) in which the starred resonances are so broad as to be unobservable, in almost the entire population. A small fraction, on the order of 0.1, remains OX and its small signals can be seen in the red spectrum. The inset at right shows the two samples, with the SQ's blue color being so intense as to render the sample almost black. Spectra were processed with 50 ms Gaussian weighting to maximize signal-to-noise and permit use of more dilute samples that facilitate quantitative reduction. NSQ, neutral semiquinone.

Flavodoxin structure, dynamics, and reactivity

Elevated temperature should accelerate exchange and cause broadening and rapprochement of the lines if they correspond to conformations exchanging slowly with one-another. However, lines broadened by intermediate exchange should sharpen as averaging accelerates at higher temperatures. Meanwhile a general sharpening is expected at higher temperatures as the decreasing viscosity of water allows more rapid molecular tumbling.

NMR spectra collected at temperatures ranging from 6 °C to 55 °C are shown in Figure 5A. Even the relatively sharp signals of Y102 and Y123 are broader at 6 °C than at 55 °C, due to slowed molecular tumbling at lower temperature. However, the resolved line of Y90a is visibly sharper at low temperature, indicating slow exchange that accelerates and contributes to the line width at higher temperatures. Indeed, at 55 °C, the widths at half-height of the Y90 signals exceeds those of Y102 or Y123. Thus, Y90 exchanges among multiple conformations, more rapidly at 55 °C (Fig. 5C). The 'a' and 'b' lines attributed to Y90 approach one-another as the temperature is raised, from 2.2 ppm apart at 6 °C to 1.6 ppm apart at 55 °C (Fig. 5D, note the extra shallow slope of Y90a's

temperature dependence and the extra steep slopes of the Y90b and Y90c resonances, as they converge). These observations confirm that Y90's multiple conformations are in slow exchange throughout the temperature we were able to examine. Coalescence temperature is above what we can explore, as protein aggregation occurs above 55 °C.

The resonance of Y121 also became broader and shifted as the temperature was raised but in a lower temperature range (Fig. 5, A and C). Indeed, Y121 broadened beyond resolution by 35 °C. The chemical shift of Y121 also became less negative at increasing temperatures, with a slope much too steep to be explained by the temperature dependence of the reference shift alone. This indicates that the unidentified signal of the other conformation is at a less negative chemical shift.

Interestingly, the resonance of Y53 is slightly but reproducibly split at 25 and 35 °C (Fig. 5), although the two conformations responsible are evidently in fast exchange at 55 °C and the general broadening due to slow tumbling impedes their resolution at 6 °C. Coalescence occurs near 45 °C. These signatures of dynamics affecting Y53, Y90, and Y121 show that three of the loops that were poorly defined in the B chain of

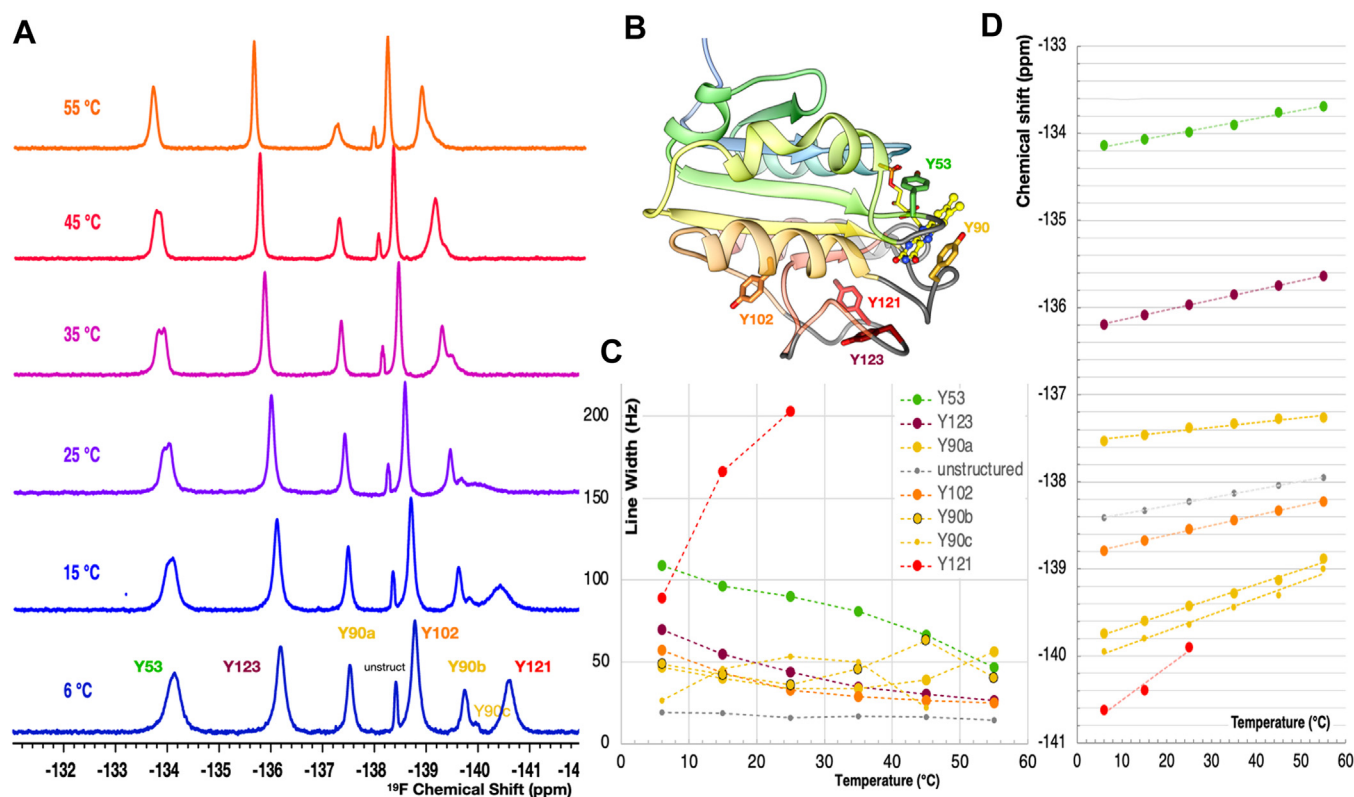


Figure 5. Temperature dependence of NMR resonances identifying dynamic regions. A, ¹⁹F-NMR spectra of ¹⁹F-Tyr Rp9Fld as a function of temperature. Resonance assignments are provided on bottom spectrum, and the corresponding residues are shown in the ribbon structure of the A chain at 100 K, in panel (B). C, comparison of temperature dependencies of line widths at half-height as a function of temperature, showing Y121's resonance sharpening at lower temperature (smaller line width), while resonances of Y102 and Y123 sharpen at elevated temperature, as for simple tumbling of a rigid protein. Increased line widths at higher temperature identify Y121 and 90 as affected by slow dynamics at RT. Y90 clearly populates ≥ 2 conformations and Y53's resonance is also seen to be split at intermediate temperatures. All three affected Ys occur in or adjacent to regions not structured in B chain at RT (Panel B). D, comparison of temperature dependence of chemical shifts. The small sharp resonance near -136.8 ppm corresponds to unstructured ¹⁹F-Tyr in unfolded protein and informs on the temperature dependence expected in the absence of conformational dynamics. The chemical shift of Y90a approaches that of Y90b as the temperature increases. Y121's chemical shift also climbs steeply, indicating existence of a cryptic resonance at less negative chemical shift with which the visible resonance will eventually merge. The temperature dependence was not extended to higher temperatures due to evidence of slow protein aggregation at 55 °C, where data were collected in blocks so that only blocks completed before onset of significant denaturation/aggregation were pooled. Spectra were processed using a 100 ms Gaussian to diminish its contribution to line widths.

the crystal structures (bearing residues 54, 86, 93–94, and 117–125) are indeed affected by dynamics. The fact that it is observed in solution exonerates crystal packing as the cause of heterogeneity. These separate stretches of the peptide chain are brought together by Fld's fold, where they surround the flavin head group and include the long loop (Fig. 2C).

Assessing possible functional perturbation of fluorinating Tyr side chains

Although ^{19}F NMR is highly perceptive of dynamics, its utility is limited if it alters Fld's functional properties. Principal among these are the reduction midpoint potentials relating the oxidized state to the NSQ state ($E^{\circ}_{\text{OX/NSQ}}$), and the NSQ state to the hydroquinone (HQ) state ($E^{\circ}_{\text{NSQ/HQ}}$). To test for perturbation of these, we employed direct electrochemistry. *Rp9Fld* was immobilized on Toray carbon paper electrodes by entrapment in a cross-linked conductive hydrogel of linear PEI functionalized with pyrene groups (pyrene-LPEI). Entrapment in a hydrogel increases the amount of protein that can be immobilized on an electrode (46), and incorporation of pyrene in the polymer improves electrical contact with the electrode, while preserving the protein's natural structure (47). The pyrene groups bind to the carbon surface *via* noncovalent pi-pi interactions (48) and the amino groups of the polymer may interact with the patch of negative surface charge with which Fld binds to its natural partners. Thus, pyrene-LPEI has been shown by Minter's group to promote direct electron transfer by redox enzymes (46).

WT *Rp9Fld*, ^{19}F -WT *Rp9Fld*, and ^{19}F -Y90W-Fld were characterized by cyclic voltammetry (CV) and square wave voltammetry (SWV) over a range of pHs (4.0–9.0). In each case, results obtained with replicate electrodes were compared. The CV of WT-*Rp9Fld* shows well-defined and reproducible

cathodic and anodic waves, as expected (Fig. 6). The peak-to-peak separation for the NSQ/HQ couple of about 58 mV confirms the reversibility of the electron acquisition/release process while the larger ΔE (98 mV) observed for the OX/NSQ couple suggests that a quasi-reversible event is coupled to the electron transfer as also reported in literature (42). We attribute this to the protonation of N5 and rotation of peptide backbone known to accompany the $\text{OX} \rightarrow \text{NSQ}$ reduction of Fld (49). Each redox couple's anodic and cathodic potential values were averaged to yield E° s, which are compared with literature values in Table S2. All values were further confirmed by SWV data.

Our value of $E^{\circ}_{\text{NSQ/HQ}} = -450 \pm 2$ mV is very close to the -434 mV obtained for *Rp8Fld* (40), but our $E^{\circ}_{\text{OX/NSQ}} = -186 \pm 4$ mV is significantly higher than the -241 mV obtained for *Rp8Fld* (E° is the E° at pH 7.0). *A. vinelandii* Fld II has $E^{\circ}_{\text{OX/NSQ}} = -187$ mV and $E^{\circ}_{\text{NSQ/HQ}} = -483$ mV when free in solution, while $E^{\circ}_{\text{OX/NSQ}} = -164$ mV and $E^{\circ}_{\text{NSQ/HQ}} = -454$ were obtained for the same Fld adsorbed to electrodes (42). The 20 to 30 mV elevation of the E° s could reflect a slight perturbation of the protein's chemical microenvironment including desolvation of the flavin and the surface chemistry of the electrode. The cationic nature of the LPEI polymer is consistent with elevation of $E^{\circ}_{\text{OX/NSQ}}$ of immobilized *Rp9Fld* versus free *Rp8Fld* (46) but is larger than expected, so we suspect that the protein itself, in addition to its state in the measurement, raises $E^{\circ}_{\text{OX/NSQ}}$ more in *Rp9Fld* than in *Rp8Fld*. Indeed, the value for *Rp9Fld* agrees much better with literature than did the *Rp8Fld* result (Table S2).

Having confirmed the method, we assessed the effect of replacing Tyr with 3- ^{19}F -Tyr. $E^{\circ}_{\text{NSQ/HQ}}$ was 26 mV less negative in fluorinated Fld than the natural version, and for $E^{\circ}_{\text{OX/NSQ}}$, the change was 8 mV. These are similar in magnitude to the 17 mV change produced by replacing Y90 with W (Fig. S3 and Table S2, also see (50)), wherein W is present naturally at that position in the LC Fld from *Methylobacter* locus PPC92239 (51), and many SC Fld and Fld domains. Thus, such variations in E° s appear to be within the functional spectrum.

Besides directly affecting redox tuning, the incorporation of 3- ^{19}F -Tyr could perturb proton equilibria, since fluorination lowers the pK_a of Tyr from 9.9 to 8.4 (52). This is not expected to alter the energetics of converting OX to NSQ, which is an electrostatically neutral event wherein the coupled proton is acquired by the flavin itself. Ionization of Y90 at high pH could affect $E^{\circ}_{\text{NSQ/HQ}}$, but Fig. S3 shows essentially superimposable pH dependence for nonfluorinated *Rp9Fld*, the ^{19}F -Y version, and ^{19}F -Y Y90W-*Rp9Fld*. Thus, our electrochemical measurements indicate no significant perturbation of Fld reactivity due to incorporation of our 3- ^{19}F -Tyr probes.

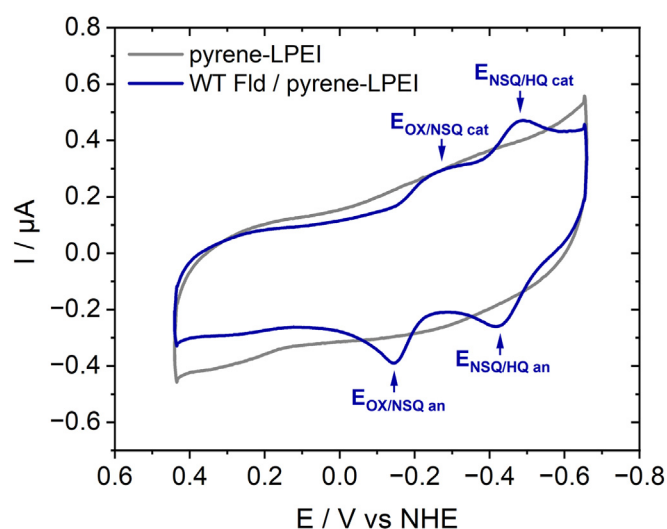


Figure 6. Direct electrochemical characterization of *Rp9Fld*'s 2 E° s. Cyclic voltammogram of WT *Rp9Fld* immobilized in pyrene-LPEI on a carbon paper electrode (blue) versus an electrode modified with only pyrene-LPEI (gray) in pH 7.0 100 mM phosphate buffer versus NHE. E_1 and E_2 indicate the potentials for the NSQ/HQ and OX/NSQ redox couples, respectively; cat and an indicate the values detected in cathodic and anodic sweeps. The scan rate was 5 mV s^{-1} . NSQ, neutral semiquinone.

A basis for dynamics in the LL's sequence?

Given that the heterogeneity evident in the crystal structure and the conformational dynamics observed by NMR are in the same region of the structure, surrounding the flavin and flanked by the LL, we asked whether the amino acid sequence

Flavodoxin structure, dynamics, and reactivity

of the LL emphasizes small residues with fewer structural constraints or residues less favored in persistent secondary structure. Moreover, given that LC Flds are proposed to employ the LL in binding to partners (36), this would subject the LL to multiple evolutionary constraints in organisms such as *R. palustris* that must use a single Fld for all partners. Therefore, we compared amino acid conservation in the LL with the conservation of amino acids elsewhere in the sequence of LC-Fld and looked for conserved flexible residues (Fig. S4).

While loops have a tendency to be more variable than core secondary structure (53), we found that the consensus sequence for LL conserved several small flexible amino acids: G, A, S, and P which together occurred 6 to 7 times (3–4, 1, 1 and 1 times, respectively) in this 21-amino acid stretch, yielding 29 to 33% prevalence compared to an average prevalence of 23% in the entire sequence and an expectation of 28% for proteins in general ((54), https://proteopedia.org/wiki/index.php/Amino_acid_composition). The elevated abundance of G, S, A, and P in the LL is consistent with enrichment expected in loops (55). In contrast, aromatics occurred four times in the LL, (2 F, 1 Y, and 1 W, 18% prevalence) *versus* a prevalence of 8.2% in the complete sequence, adjusted to an expectation of 6.1% for loops (Fig. 7A).

Long loop interactions with Fld's partners

The LL contains several conserved aromatic residues and an anionic motif (Fig. 7A). The Y121 we identified with an exceptionally broad resonance is almost fully conserved (96.0% as Y) and the next most abundant residue at that position is F (3.1%) followed by H (0.9%) demonstrating conservation of an aromatic group smaller than Trp. The position of Y123 is most frequently held by F (77.2%) with Y next (12.9%) and then His (7.6%) (Also see Fig. S4 for a color-coded multiple sequence alignment (MSA)). Y123 is followed by a conserved pattern of 3 to 5 D/E/S residues that forms an anionic/polar stretch. These add to a prominent halo of negative charge surrounding the site at which the edge of the flavin protrudes into solvent between Y53 and Y90 (Fig. S1). Thus the LL contributes to an extensive anionic interaction surface mimicking that of Fd (37, 56–58) and presumably aiding in binding to Fd's partners.

For a counter-test of LL contributions to Fld's affinity and selectivity for partner proteins, we extended our sequence comparison to Flds that do not need to identify and capture their partners: Fld sequences that occur within a larger protein (consisting of an Fld domain fused to a second domain). Selectivity is essentially assured by the linker uniting the two domains, and the local concentration of each domain is so high from the perspective of the other that a relatively weak

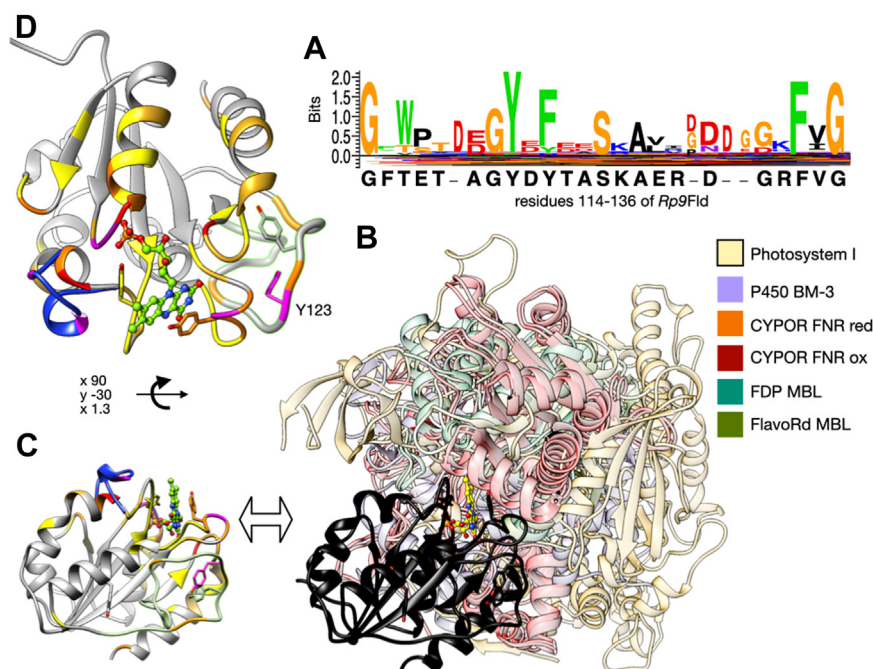


Figure 7. Sequence conservation in the LL and Fld residues that interact with partner domains. A, sequence Logo for the long loop of Rp9Fld showing the relative abundances of the different amino acids found at each position and comparing the consensus identity with the identity in Rp9Fld's sequence. Panel B, overlay of complexes of Fld with a partner protein or partner domain. All complexes were overlayed on Rp9Fld using their Fld domain only; however, only the other domains, with which it interacts, are shown. These are as follows: *wheat* = complex with photosystem I based on cryoEM (6KIF (37)); *mauve* = cytochrome P450 domain of cytochrome P450-BM3 (1BVY (89)); *red* = the Fd NADP⁺ reductase (FNR) domain of closed conformation NADPH cytochrome p450 Oxidoreductase (CYPOR) with NADP⁺ bound (5URE (67)); *burgundy* = oxidized CYPOR FNR domain (1AMO (101)); *blue-green* = FprA, nitric oxide reductase's metallo-β-lactamase (MBL) domain (1YCH (45)); *olive green* = MBL domain of flavorubredoxin, another flavodiiron protein (4D02 (44)). The Rp9Fld template is shown in *black* and its FMN is in *yellow*. See Fig. S5 for compositions of the modular proteins. Panel (C) applies color-coding to the same orientation of Fld to identify residues found to interact with a partner protein in any of the complexes or based on NMR chemical shift changes upon exposure to partner in solution (12, 24). *Gray* = non-interacting; *yellow, amber, orange, red, pink, and purple* indicated increasing frequency of interaction, see Experimental procedures for details. A loop specific to Fld associated with nitrogenase is shown in *blue*. Panel (D) shows a second orientation of Rp9Fld, from the partners' perspective, facing towards the protruding flavin. In (C and D), Rp9Fld's flavin is in *green*, the LL is highlighted in *lime green* and thickened in (D) to facilitate reference to the Logo in panel (A). LL, long loop.

interaction can suffice to bring the two together for long enough for electron transfer. Therefore, one expects that evolutionary pressure to optimize selectivity and affinity by other means would be diminished, and any residues in LL selected for these functions in independent Fld would be less strongly conserved in Fld domains.

We analyzed 147 members of the larger protein family identified with *Rp9Fld*'s amino acid sequence: IPR008254. Co-alignment of the sequences of their Fld domains with the above set of 222 LC-Flds and 103 SC-Flds reveals that the Fld domains possessed by members of IPR008254 all lack an LL (Fig. S5). This is negative information but nonetheless suggests that the selectable benefits of an LL are not needed by attached Fld domains. This is consistent with a role for LL in partner binding and selection, two challenges unique to independent Flds.

Since LL appears to occur only in independent Fld, whose complexes with partners tend to be short-lived (57, 59), there is relatively little structural information on the placement of LL within complexes. However, the Fld residues in contact with partner were identified by a cryo-EM structure of a cryo-trapped complex of LC flavodoxin with photosystem I (37) and were implied by NMR resonance shifts upon weak association with nitrogenase in solution for *Azotobacter chroococcum* Fld ('NifF') (12) and more extensive resonance shifts observed for *Escherichia coli* Fld's associations with FNR or methionine synthase (24). Figure 7B combines these with additional insights from interactions formed by Fld domains in modular proteins, for which crystal structures have been solved. Based on eight studies, each Fld residue was color-coded by the number of complexes in which it interacts with partner. Panels 7 C and D show that the partner proteins bind to a shared surface of the Fld protein, with the participant residues contributed mostly by loops but also subtended by helices 1 and 5. These residues surround the site of flavin exposure (Fig. 7D).

Discussion

Fld has dynamic loops, near the flavin

NMR and crystallography show that the LC Fld of *R. palustris* CGA009 explores considerable conformational variation in the region that surrounds the isoalloxazine ring. Residues in these surface loops adopt distinct conformations in the two protomers in our crystals' asymmetric unit; they are unresolved in the B protomer, and ordering was dependent on temperature, consistent with dynamic disorder. Although residues in the loop containing Y53 and conserved G54 and G56 are known to accommodate redox-coupled rotation of the backbone at G54 that reverses a H-bond with the flavin N5 (60), our X-ray data collected at 100 K reveal heterogeneity in the LL and the loop containing Y90 as well. The divergence between the two protomers in our unit cell documents that these loops can adapt to different partner protein interfaces. This has previously only been observed for apo-FMN Flds (61). However, in our case, the FMN density is well resolved.

NMR concurs, providing direct evidence of conformational exchange affecting Tyr side chains in the same loops. We also confirm the utility of ^{19}F NMR in conjunction with the

incorporation of 3- ^{19}F -Tyr for characterizing these dynamics without disrupting the functional redox reactions of *Rp9Fld*. Having assigned the resonances of all five Tyr and demonstrated retention of LL signals even in the NSQ state, we are now equipped to use ^{19}F NMR to monitor side chains interactions with partner proteins. The LL's conserved anionic motif and its retention in many independent Flds but not Fld domains that are fused to their partners suggest evolutionary pressure on the LL specifically when affinity and selectivity for a partner are needed.

Loop dynamics are clarified by comparison of Fld from different strains of *R. palustris*

Although *Rp9Fld* is 75% identical to the *Rp8Fld* we characterized before (40), comparison of the two resolves some puzzling behaviors seen in *Rp8Fld* and provides a more nuanced view of the LL. Three of the four Tyr shared by *Rp9Fd* and *Rp8Fld* retain qualitatively the same temperature dependencies in the two Flds, with Y90 in slow exchange but approaching intermediate at 55 °C, Y121 in intermediate exchange, and Y102 in fast exchange, consistent with Y102's surface-exposed side chain that experiences a relatively constant environment. (We rule out occupation of a single conformation because 3-F-Tyr bears a single ^{19}F that would occupy one of two positions if the Tyr ring were immobile.) The lower coalescence temperature for Y121 than for Y90 need not mean a different rate constant for conformational change, as the separation between peaks characterizing the participating conformations also governs the rate sufficient to average the resonance frequencies. The lower temperature coalescence of Y121 more likely indicates that the slow exchange Y121 resonances are closer together than those of Y90, and thus that the environments sampled are less different for Y121.

In *Rp8Fld*, the two half-strength signals of Y90 could have reflected the two orientations of the Tyr ring. However, the three signals we observe in *Rp9Fld* can no longer be explained by slow flipping of the Y90 ring alone and indicate that Y90 dwells in at least two different well-defined environments, besides the two sampled *via* ring flips. This could result from repositioning of the LL since it is very near Y90 (Fig. 2C), although the 90s loop also displayed backbone dynamics (62). Given that Y90's signals approach intermediate exchange in a physiologically relevant temperature range, we speculate that transient unstacking of Y90 from the flavin may allow accelerated electron transfer and that such unmasking of the flavin could be rendered responsive to partner binding by interactions of the LL with both the partner and the 90s loop.

In *Rp8Fld*, only one of Y90's resonances appeared sensitive to flavin paramagnetism, raising big questions as to the nature of the conformation in which Y90's side chain appeared distant from the flavin (40). *Rp9Fld*'s behavior is in much simpler agreement with expectations, as formation of the semiquinone state uniformly and almost quantitatively obliterates resonances of all signals of both the two Tyr that flank the flavin, without affecting the other Tyr resonances. Thus, *Rp9Fld* emerges as a case in which paramagnetism can be a highly

selective means of identifying residues that come close to the flavin. This system will provide a promising tool for mapping out the Fld binding footprint on partner proteins (63).

The improved spectra of *Rp9*Fld also reveal a subtle but unambiguous splitting of Y53's resonance consistent with slow exchange dynamics in the 50s loop that has been documented in NMR relaxation studies of backbone (39, 62) but was not seen in the ^{19}F NMR signals of side chains. This signature of multiple conformations is particularly interesting given that Y53's side chain H-bonds with the phosphate of FMN, while the backbone of the 50s loop provides redox-coupled H-bonding to the flavin N5 from the amide of residue 55 when the flavin is oxidized but to the carbonyl of G54 when the flavin is reduced (60).

Finally, we get richer insight into dynamics in the LL, thanks to presence of a second probe in the form of Y123. While Y121's side chain is in intermediate exchange at 25 °C, that of Y123 appears to be in fast exchange (see regarding Y102, above). This is consistent with our crystal structures based on greater constraint of the Y121 side chain and more different environments on the two sides of Y121's ring. Based on the A chain in our structure, the side chain of Y121 is more buried than that of Y123 (side chain SASA of 8.95 Å² for Y121 *versus* 13.46 Å² for the side chain of Y123), suggesting that side chain rotation will be slower and chemical shift averaging less efficient.

Interactions with partners

FNR (Fd-NADP⁺ reductase) uses electrons from NADPH to reduce Fd or LC Fld; reduced Fld in turn reduces methionine synthase (24) (Fig. S5). Changes in NMR chemical shifts identified Fld residues affected by binding to each of these partners, showing that both interact *via* the face from where Fld's flavin protrudes, including the loops that surround the flavin (Fig. 7B) (24). The two partners' footprints are not identical, but they overlap considerably (24).

E. coli's LC Fld also reduces pyruvate formate lyase activating enzyme but only binds to it when FMN-replete (64). Since acquisition of FMN rigidifies Fld's core and loops surrounding the FMN (39, 61), this indicates that partner binding requires some preformed loop structure and implicates the protruding flavin itself as a contributor.

Fld also supplies low- E° electrons to nitrogenase, and titration of *Azotobacter croococcum*'s LC Fld with V-utilizing nitrogenase's Fe protein identified three responsive residues *via* chemical shift changes (12). All surround the flavin, but two fall within the loop that is absent from our crystal structure (blue in Fig. 7, B and C). Indeed, our MSA reveals that the longer loop noted in *A. vinelandii* and *R. capsulatus* is associated specifically with organisms that fix N₂ (18) (black box in Fig. S4). This supports a role for surface loops as agents of partner selectivity.

Independent Fld may be evolved to not form long-lived associations with partners precisely because slow detachment from a source partner would also diminish the rate at which acquired electrons could be delivered to their destination

partner. Besides the cryotrapping employed as part of cryo-EM (37), connections between Fld domains in modular proteins serve as natural cross-links enabling crystal structures of complexes. These all involve SC Flds (8) but nevertheless reveal that much of the surface on which the LL lies is not used to interface with partners. By masking this surface, the LL could 'steer' partners to productive surfaces instead and/or it could enlarge the binding surface. To imagine how the LL might augment a productive interaction or impede an unproductive one, we looked at interactions of Fld domains in their larger multidomain proteins.

A prevalent combination is fusion of a domain related to FNR (containing FAD) behind an Fld domain (containing FMN) in the so-called diflavin reductases (65, 66) exemplified by the NADPH:cytochrome p450 oxido reductases (CYPORs, Fig. S5) (67). CYPOR's "closed" conformation brings the FNR and Fld domains together to favor electron transfer from FAD to FMN (red and burgundy in Figs. 7B and S6); however, an "open" form is required to enable FMN to reduce the P450 heme (68). In the self-sufficient P450-BM3, the cytochrome p450 domain is appended to the N terminus of the CYPOR to produce a modular enzyme with the domain sequence n-P450-Fld-FNR-c (69). Removal of the FNR domain enabled determination of a crystal structure modeling Fld-P450 interactions (purple in Fig. 7B, (69)). (Note that in dimers of P450-BM3, the Fld domain of one monomer may reduce the cytochrome of the other (70).) Similarly, in the paralogous human methionine synthase reductase, the Fld domain reduces a corrin-containing methionine synthase (24) (Fig. S5). These three representative structures position the flavin of the FNR domain or the heme of the P450 domain close to the flavin of Fld (Fig. S6).

Flavodiiron proteins are a different modular combination that fuses a Fld domain to a preceding metallo- β -lactamase-like (MBL) domain, where a diiron site metabolizes O₂ or NO (71, 72) (Fig. S5). The interface between the Fld and MBL domains is seen to occur between protomers of the dimer, associating the Fld domain of one protomer with the MBL domain of the other. Figure 7B also includes the complex of Fld with photosystem I (6KIF). All are overlaid by superimposing their bound Fld domains, and Figure 7, B–D show that the different partners all bind to the face of Fld from where the flavin protrudes. Additionally, the partners all present their own redox cofactor at the surface or only slightly buried, so that the donor and acceptor cofactors are brought together (Fig. S6).

Regarding the LL, it is striking that only the early portion of the LL, bearing Y123 and Y1221, interacts with partners. That portion is near the flavin and increases the size of the binding interface but does not seem to justify conservation of such a large loop. We speculate that the later portion of the LL may steer Fld into productive orientations by masking a large portion of the Fld surface whose association with partner would place Fld's flavin further away from partner cofactors. Indeed, the LL would impede formation of the dimers observed in the SC Flds of *Methanosarcina acetivorans* and *Dosidicus gigas* (4HEQ and 5WID, Fig. S2). Meanwhile, the

early portion of the LL contributes to the anionic halo believed to attract Fld to partners (37, 57, 58).

Given the conservation of aromatic residues in the LL, we wondered if they could play a role in electron transfer, in effect providing a conductive interface (57, 64, 73). If so, they might be enriched in the regions that interface with partners. The 56 residues found to interact with partners includes five aromatics (Fig. S7A), compared to expectation of 4.6 based on the prevalence of Y, F, and W overall (https://proteopedia.org/wiki/index.php/Amino_acid_composition) and 3.4 after correction for their expected prevalence in loops (55). Thus, the level of conservation we observe on these surface loops (Fig. S7B) supports a role for aromaticity, which we propose to be mediation of electron transfer. By extension, the aromatics at positions 121 and 123 of the LL may do the same. They are folded inward in our crystal structure, but we propose that the dynamics we have documented enables them to present themselves in the interface upon partner binding.

Concluding remarks

Although the LL is structured in the A protomer of our crystal structure, our NMR evidence indicates that it and the two loops containing Y53 and Y90 could adopt different conformations as part of binding with partner protein. All three are exposed on surfaces that interface with partners of Fld, and Y53 and Y90 protect the flavin from solvent in the absence of partner but could unmask it upon partner binding. Indeed, the LL has two conserved motifs that argue for a contribution to function. We speculate that the LL contributes to binding affinity by enlarging the binding interface, while also buttressing Y90. Interactions with partner could reposition the LL, effectively freeing Y90 to interact with the partner and unmasking the flavin. With the ^{19}F -Tyr resonances assigned, we look forward to titrating *Rp9Fld* with its candidate partners, to learn whether Y121 and Y123 are indeed responsive, interacting with partners and possibly even enhancing electron transfer.

Experimental procedures

Purification of *Rp9Fld* and incorporation of ^{19}F -Tyr

The expression system employed a synthetic gene for *Rp9Fld* (GenScript) using codons optimized for *E. coli* and the amino acid sequence of flavodoxin from *R. palustris* strain CGA009 (*Rpa9Fld*). It was provided in the commercial vector pET28b(+). Primers encoding four of the five desired amino acid substitutions were designed using NEB Base Changer software (See Table S1) and ordered from Eurofinn Genomics. They were employed along with the WT-Fld bearing parent plasmid as template in PCR utilizing Q5 High Fidelity DNA polymerase from New England Biolabs. The fifth mutant gene (encoding the Y90W substitution) was synthesized and cloned into pET28b(+) by GenScript. The WT and mutant plasmids were transformed into competent NiCo21(DE3) *E. coli* cells for protein production (New England Biolabs C2529H) and NEB 10-beta competent *E. coli* (High Efficiency, Catalog# C3019H) for plasmid propagation.

For protein expression, an initial 10 ml overnight culture of plasmid-bearing *E. coli* strain Nico21DE3 was incubated at 37 °C and used to inoculate 1 l of terrific broth (TB) medium containing 50 µg/ml and 0.09 µM kanamycin for selection. The culture was agitated at 250 RPM at 37 °C until the optical density reached the range of 0.6 to 0.8. The temperature was then lowered to 21 °C, and protein expression was induced by the addition of IPTG to 0.5 mM. Following 18 h of post-induction incubation, the cells were harvested by centrifugation, resuspended in PBS pH (137 mM NaCl, 2.7 mM KCl, 10 mM K_2PO_4 , and 1.8 mM KPO_4) and pooled before pelleting again for preservation at -80 °C until use.

To obtain purified protein, frozen cells were thawed and suspended in a binding buffer consisting of 20 mM Tris-HCl, 300 mM KCl, and 10 mM imidazole at pH 7.4, which was augmented with FMN monosodium salt (1 mM, Cayman chemicals, catalog #18167), Benzonase nuclease (1 µl from Millipore sigma # 71205-3), and recombinant lysozyme (1 µl, Millipore sigma #: 71110-4). Cell lysis was achieved through sonication, employing a pulse mode with 5 s on and 30 s off for a total of 30 cycles, while maintaining a cold environment with ice. The resulting cell lysate was loaded onto a 5 ml His-Trap column (Cytiva, Histrap HP, catalog: 17524802), pre-equilibrated with 20 mM Tris and 300 mM KCl at pH 7.5, using the ÄKTA GO FPLC system. The column was washed with 20 column volumes of the binding buffer supplemented with 20 mM imidazole. Finally, the target protein was eluted from the column using 250 mM imidazole in 20 mM Tris-HCl, 300 mM KCl, plus 250 mM imidazole, pH 7.4. The eluted protein was concentrated and subjected to buffer exchange by gel filtration over a DG-10 column equilibrated with 20 mM bis-tris propane buffer with 200 mM KCl at pH 7.5. The His₆ tag was not removed. A typical yield from 1 l of culture was 48 mg protein from a 1 l culture.

Fluorination for NMR experiments involved use of initial growth in TB, before transfer of the cells into defined medium (74) withholding Tyr and providing Phe, Trp, and ^{19}F -Tyr, as described by Varner *et al.* (40). Preparation of fluorinated protein was also as per Varner *et al.* and above, except that sonication was used to lyse cells. Thus we obtained a purified protein yield of 22 mg of ^{19}F -WT-Fld from a 1 l initial culture of TB medium.

Flavin content and protein concentration were determined as per literature (40) and using the Pierce 660 nm method (# 22660) versus bovine serum albumin standard, respectively (75). The flavin stoichiometry was determined to be 1.15 ± 0.3 FMN per *Rp9Fld* monomer, and optical spectra of the WT, the fluorinated WT, and variants employed for signal assignment are provided as Fig. S8. To optimize purity of the protein, for crystallography, we employed gel filtration chromatography on a Cytiva Superdex 200 Increase 10/300 G1 column (column pressure: 2 Mpa, flow rate: 0.5 ml/min, volume of each elution fraction: 0.5 ml). Fractions corresponding to the protein peak were pooled and concentrated, prior to use in crystallization experiments. Gel filtration chromatography was used as above, in conjunction with a mass-calibration (Biorad gel filtration standard, # 1511901) to determine the effective mass of

Flavodoxin structure, dynamics, and reactivity

Rp9Fld. A single symmetric elution band was obtained yielding a value of 17.5 kDa $\pm$ 5% for *Rp9Fld* concentrations up to 1 mM. Thus, *Rp9Fld* is >95% monomeric in solution at the 0.55 mM concentration used for SEC.

Crystallization of *Rp9Fld*

Rp9Fld was crystallized using the hanging drop vapor diffusion method using 8 to 12 mg/ml protein stock solution in 20 mM bis-tris propane buffer, pH 7.5, with 200 mM KCl. Preliminary success was achieved using the Hampton screen (HR2-139) within 96-well sitting vapor diffusion setups. We then refined the concentrations of magnesium formate (0.2 M) and PEG 3350 (20%), along with the addition of an additive (Hampton Research HR2-464), and varied the temperature (4–22 °C). The mother liquor contained 0.3 to 0.6 M magnesium formate with 15 to 20% PEG 3350. Flower-shaped crystal clusters emerged after 2 weeks at 14 °C. These were utilized to generate 'seed-crystals' using Hampton Research's Seed Bead Kit (HR2-320). Diffraction quality and crystal morphology were optimized by iterative cycles of micro seeding and screening of crystallization conditions. Several morphologies of crystals were obtained and checked to identify crystals with the best diffraction quality in combination with the lowest mosaicity. This process resulted in the final use of 0.4 M magnesium formate pH 7.5 and 19% PEG 3350, wherein we note that the best crystal morphology was obtained upon supplementation with cadmium chloride at 20 mM. Final crystallization drops contained 1 μ l of protein solution and 0.8 μ l of 2 $\times$ crystallization solution with 0.2 μ l of micro seed solution (1:1000 dilution).

When data were collected in capillaries at room temperature, radiation damage was observed in the form of decreased $I/\sigma(I)$ and elevation of R_{sym} . We were nevertheless able to collect a room temperature data set including diffraction to 2.9 Å resolution. To obtain a deposition-quality structure, single plate-like crystals were exposed briefly to 35% W/V PEG 3350 in mother liquor solution to provide cryoprotection, before exposure to a stream of liquid N₂. Data were then collected at 100 K, yielding diffraction to 2.1 Å.

X-ray data collection and structure refinement

X-ray data were collected by using a Rigaku HighFlux HomeLab source equipped with a MicroMax-007 HF X-ray generator and Osmic VariMax optics. Diffraction images were captured using an EIGER R 4M hybrid photon-counting detector. The data were integrated using the CrysAlisPro software suite (Rigaku) and then reduced and scaled using AIMLESS (76) from the CCP4 suite (77). Initial phases for each structure were obtained by molecular replacement using 1CZN (flavodoxin from *Anacystis nidulans*, 43% identical, 98% coverage) as a starting model in Phaser (78) in the Phenix suite (79). The model was iteratively refined including manual model building using the Phenix refinement module (80) and the Coot molecular graphic program (81), respectively. Solvent atoms were added at last step of model building and refinement cycle. These were added only when reasonable 2Fo-Fc

density and at least one hydrogen bond to a donor was present. Occupancies and thermal parameters of the solvent were refined in alternating cycles.

The final model was checked and validated with Molprobrity (82). Buried surface area and solvent accessible area calculations were done using PISA software (43) and the EPPIC (Evolutionary Protein-Protein Interface Classifier) server (83). High resolution images were generated using Pymol (<http://www.pymol.org>). Statistics of diffraction data processing and the converged refinement statistics for datasets along with R , R_{free} , and other validation measures are given in Table 1, and the final structure coordinates and structure factor amplitudes were deposited in the RCSB Protein Data Bank with the accession code 8SNZ. For completeness, the structure obtained from data collected at room temperature was also deposited, as 8V2Y.

NMR spectroscopy

NMR samples were prepared in a 20 mM bis-tris propane buffer supplemented with 200 mM KCl at a pH of 7.5. 10 to 15% v/v D₂O was present to provide a frequency lock signal for instrumental stability. To permit quantitative formation of the NSQ state of *Rp9Fld*, a 2 mM Fld sample was equilibrated in an inert atmosphere of N₂ inside a glove box (Bell Technologies). After stepwise reduction with dithionite to produce the neutral blue semiquinone (inset, Fig. 4), the sample was carefully transferred to an NMR tube, which was then sealed with a septum stopper before removal from the glove box. Immediately upon removal, the septum was secured with parafilm and the sample was transferred to the spectrometer. Data collection was complete in 2 h after which the sample remained blue. Only several days later was a color change to green/yellow visible at the top of the tube. Thus, the sample remained fully in the NSQ state during NMR data collection.

¹⁹F NMR spectra were recorded using a Varian Unity/Inova 400 MHz spectrometer, at a constant room temperature of 25 °C or a Bruker NEO 400 MHz spectrometer with SMART probe. 30 ppm wide scans centered at −134.7 ppm were collected following a 10 μ s 90° excitation pulse with data collected for 0.27 s followed by a 1.5 s delay for recovery between scans. Spectra shown represent averaging of 1024 to 4096 scans, for samples of 0.4 to 2 mM *Rp9Fld*. The above recovery delay was based on our measured T_1 relaxation times, of which the longest was 620 ms. Data were processed using VnmrJ, Topspin, and Mnova software, depending on the data source and the analysis to follow. Baseline artifacts from distortion of early time points were eliminated by removing the first five data points and replacing them with backward linear prediction based on 256 complex points that followed, modeling the FID in terms of 32 coefficients. Gaussian line broadening dropping to 50% at 100 ms was employed to optimize signal to noise without broadening out reproducible splitting (See figure captions). Baseline correction was repeated multiple times and resulting peak shapes and integrals compared to identify values that were robust to processing parameters. Moreover, the spectra collected as a function of

temperature were also simulated by Lorentzian deconvolutions (VnmrJ) to yield a self-consistent set of peak areas, positions, and widths, for comparison with those measured directly from the spectra.

Bioelectrode preparation and electrochemical experiments

Ethylene glycol diglycidyl ether (EGDGE) was purchased from Polysciences Inc. All other chemicals were purchased from Sigma Aldrich and used as received without further purification. All water used in bioelectrochemical experiments was filtered using an Ultrapure MilliQ system. Pyrene-LPEI was previously synthesized using the reported procedure (46). Stock solutions of protein (WT *Rp9Fld* Fld, ^{19}F -Y WT *Rp9Fld*, and ^{19}F -Y Y90W-*Rp9Fld*) 10 mg mL⁻¹, pyrene-LPEI 10 mg mL⁻¹, and 13 vol% EGDGE were prepared in water. The crosslinked hydrogel was prepared by gently mixing 70 μL of pyrene-LPEI, 30 μL of protein, and finally 4.5 μL of EGDGE. Toray carbon paper electrodes were cut into 2 cm $\times$ 0.5 cm stripes and dipped into melted paraffin wax that was successively cooled down and solidified to delimit an electroactive surface area of 0.25 cm² (0.5 cm $\times$ 0.5 cm) at one end. A total volume of 30 μL of the polymeric mixture containing each protein were drop-coated on the exposed electroactive area of the electrodes; three successive depositions of 10 μL aliquots of the mixture were carried out followed by 20 to 30 min of drying to allow the crosslinking process to occur. Control experiments containing pyrene-LPEI without protein were made by mixing the same amount of polymer and EGDGE with water and depositing the same amount of mixture volume on the surface of the electrodes. After the last deposition, the electrodes were cured under anoxic conditions overnight to allow stabilization of the polymer matrix. Preparation of the electrodes as described as well as electrochemical measurements have been carried out inside of a glove box under an inert atmosphere of Ar/H₂ (2.8–3.2%) with less than 3 ppm O₂ at room temperature. Bioelectrodes have been tested by CV and SWV with a CH Instruments 1033 potentiostat. CVs were performed at 5 mV s⁻¹. A three-electrode set up was used with a SCE reference electrode, a platinum mesh counter electrode, and Toray carbon paper working electrodes. The waxed end was used as an electrochemical connection point for an alligator clip of stainless-steel wire. Electrodes were tested in 5 ml of sodium phosphate buffer (100 mM) at different pHs. Potentials have been recalculated with respect to a normal hydrogen electrode reference electrode. Data were elaborated with Origin 2023b software (www.originlab.com).

Amino acid sequence analysis

To compare the degree of conservation in the LL with overall conservation in Fld, we constructed an MSA of LC Fld sequences from uniref50 and compared it with an analogous MSA of uniref50's SC Fld sequences, because only the former will contain the LL. LC-Fld sequences were gathered using EFI (84, 85) and an AST of 35 based on the interpro family

IPR010086, encompassing *Rp9Fld* and *NifF* from *R. capsulatus* 2WC1 (18). SC-Fld sequences were similarly obtained based on IPR010087, the interpro family containing the Flds of *Desulfovibrio vulgaris* (3FX2, 1J8Q), *Megasphaera elsdenii* (2FZ5 (86)), and *Lactobacillus* (3EDO). Sequences were gathered from the Ref50 database, where sequences are grouped into nodes uniting sequences at least 50% identical over at least 80% of the longest sequence in the node (the so-called 'seed sequence' of that node). Nodes from the ref50 protein database were used to limit the number of sequences considered without altering the structure of the relationships among them (85).[†] In each of LC- and SC-Flds, MSAs were curated to eliminate a few sequences that were incomplete or contained an insert affecting 1% or fewer of the sequences. The separately aligned LC- and SC-Fld sequence sets were combined and coaligned to localize the LL and highlight shared motifs. After culling three more LC-Fld sequences based on this expanded perspective, the LC-Fld sequence set consisted of 222 sequences (trimmed from 237) and the set of SC-Flds comprised 103 sequences (trimmed from 106). After co-alignment to obtain a single numbering system applicable to both, the two groups were separated and the quality of residue conservation within each was compared, without further modification of the MSAs (Fig. S5).

Sequence Logos were generated by Seq2Logo (87) using the p-weighted Kullback-Leibler method without clustering, based on our alignment of 222 LC Flds. For Fig. S8, the Logo is based on our 222 aligned LC Flds but includes all positions documented to interact with partners for SC Flds and Fld domains as well (8 complexes).

To learn what aspects of Fld could be most important for orienting Fld *versus* selection or affinity for appropriate partners, we compared the sequences of independent Flds with Fld domains that are fused to a partner and thus assured of a locally high concentration. We used the Pfam family PF00258, which subsumes the two interpro families listed above and additionally incorporates multidomain enzymes in which a domain has high homology with Fld. Pfam family PF00258 is represented by 274 reviewed sequences (75 k total sequences as of Aug 2023 which have been incorporated into IPR008254 representing 444 reviewed sequences, 133 k total as of Aug 2023). We selected enzymes containing an Fld domain plus an additional domain using sequence length cut-offs of 350 to 450 with an alignment score threshold of 104. An MSA of these sequences along with those of LC and SC Flds caused the larger modular proteins to all align on the basis of their Fld domains. For the modular proteins, we deleted sequence stretches representing the other domains and advanced using the portion of each sequence that represents the Fld domain. Thus we collected 156 Fld domain sequences which we culled to 147 ranging from 144 to 202 residues long. To permit direct

[†] LC-Fld IPR010086 had 9,269 seq in Uniprot, 2894 in Uniref90 and 314 in Uniref50 (fragments were excluded) as of July 2023. SC-Fld IPR010087 had 4827 sequences in Uniprot, 1766 in Uniref90 and 136 IDs in Uniref50, as of July 2023.

comparison with available crystal structures, we augmented this set with the Fld domain sequences corresponding to pdb files 1BVY, 1E5D, 4H2D, 1YKG, and 1AMO.

To map residues involved in partner binding onto the structure of R_p9Fld, residues were colored according to the frequency with which they were reported to interact with partner protein or domain, in structures of complexes with partners, where possible. Color coding indicates the frequency with which each residue interacts with partner, based on crystal structures of complexes listed in the caption, changes in NMR chemical shifts upon titration with partner (12, 24), or cryo-EM structure of complex (37). For each instance in which a residue interacted with partner (distance of 3.5 Å or less) or listed as interacting by author, the residue acquired a 'point' since there were five structures of SC Fld but only three of LC Fld; instances observed for LL residues were possible in only ≈ one-third of cases, so they were accorded 3 points each. The sum of points obtained by each residue which then determined its color as follows: 0 = gray, 1 = yellow, 2=gold, 3 = orange, 4 = red, 5 or 6 = pink, 7 = purple. Information was only available for one diazotroph, so responsive residues in the loop unique to that case were shown in pink and that loop is in blue.

Data availability

Data are either in this manuscript, in the [Supporting information](#) or deposited in the PDB as 8SNZ and 8V2Y.

Supporting information—This article contains supporting information (16, 24, 37, 40, 42, 61, 65–67, 69–71, 88–99).

Acknowledgments—The authors would like to thank Drs Dylan Boucher and Dean Myles for helpful discussions and Messrs. A. Sebesta and T. Stone for infrastructural support. A portion of this research at ORNL's Spallation Neutron Source facility was sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy.

Author contributions—S. K. and A. A. methodology; S. K., A. A., M. B., and A.-F. M. investigation; S. K., D. D., S. M., and A.-F. M. resources; S. K., A. A., M. B., and A.-F. M. validation; S. K., A. A., M. B., and A.-F. M. visualization; S. K., A. A., M. B., D. D., W. E. H., S. M., and A.-F. M. writing—review and editing; A. A. and A.-F. M. writing—original draft; S. M. and A.-F. M. conceptualization; S. M. and A.-F. M. supervision; S. M. and A.-F. M. funding acquisition.

Funding and additional information—This work was supported by NSF grant number CHE-CMI 2154206 to S. M. DOE DE-SC0021283 and KY-EPSCoR PON2 635 2000003148 funding to A. F. M. NSF CHE 2108134 supported contributions of D. D. to sample production.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CV, cyclic voltammetry; CYPOR, cytochrome p450 oxidoreductase; HQ, hydroquinone; LC, long chain; LL, long loop; MBL, metallo-β-lactamase-like; MSA,

multiple sequence alignment; NSQ, neutral semiquinone; SC, short chain; SWV, square wave voltammetry; TB, terrific broth.

References

- Warburg, O., and Christian, W. (1932) On a new oxidation enzyme (Über das neue oxydationsferment). *Naturwissenschaften* **20**, 980–981
- Blyth, A. W. (1879) The composition of cow's milk in health and disease. *J. Chem. Soc. Perkin. Trans.* **35**, 530–539
- Michaelis, L., Schubert, P., and Smythe, C. V. (1936) Potentiometric study of the flavins. *J. Biol. Chem.* **116**, 587–607
- Knight, E., Jr., D'Eustachio, A. J., and Hardy, R. W. (1966) Flavodoxin: a flavoprotein with ferredoxin activity from *Clostridium pasteurianum*. *Biochim. Biophys. Acta* **113**, 626–628
- Knight, E., Jr., and Hardy, R. W. F. (1966) Isolation and characteristics of flavodoxin from nitrogen-fixing *Clostridium pasteurianum*. *J. Biol. Chem.* **241**, 2752–2756
- Cotruvo, J. A., Jr., and Stubbe, J. (2008) NrdI, a flavodoxin involved in maintenance of the diferric-tyrosyl radical cofactor in *Escherichia coli* class Ib ribonucleotide reductase. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14383–14388
- Mayhew, S. G., and Tollin, G. (1992) General properties of flavodoxins. In: Müller, F., ed. *Chemistry and Biochemistry of Flavoenzymes*, CRC Press, Boca Raton, FL: 389–426
- Karlusich, J. J. P., Lodeyro, A. F., and Carrillo, N. (2014) The long goodbye: the rise and fall of flavodoxin during plant evolution. *J. Exp. Bot.* **65**, 5161–5178
- Fixen, K. R., Chowdhury, N. P., Martinez-Perez, M., Poudel, S., Boyd, E. S., and Harwood, C. S. (2018) The path of electron transfer to nitrogenase in a phototrophic alpha-proteobacterium. *Environ. Microbiol.* **20**, 2500–2508
- Campbell, I. J., Bennett, G. N., and Silberg, J. J. (2019) Evolutionary relationships between low potential ferredoxin and flavodoxin electron carriers. *Front. Energy Res.* **7**, 79
- Tognetti, V. B., Zurbriggen, M. D., Morandi, E. N., Fillat, M. F., and Valle, E. M. (2007) Enhanced plant tolerance to iron starvation by functional substitution of chloroplast ferredoxin with a bacterial flavodoxin. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11495–11500
- Peelen, S., Wijmenga, S., Erbel, P. J., Robson, R. L., Eady, R. R., and Vervoort, J. (1996) Possible role of a short extra loop of the long-chain flavodoxin from *Azotobacter chroococcum* in electron transfer to nitrogenase: complete 1H, 15N and 13C back-bone assignments and secondary solution structure of the flavodoxin. *J. Biomol. NMR* **7**, 315–330
- Lodeyro, A. F., Ceccoli, R. D., Karlusich, J. J. P., and Carrillo, N. (2012) The importance of flavodoxin for environmental stress tolerance in photosynthetic microorganisms and transgenic plants. Mechanism, evolution and biotechnological potential. *FEBS Lett.* **586**, 2917–2924
- Bianchi, V., Eliasson, R., Fontecave, M., Mulliez, E., Hoover, D. M., Matthews, R. G., et al. (1993) Flavodoxin is required for the activation of the anaerobic ribonucleotide reductase. *Biochem. Biophys. Res. Commun.* **197**, 792–797
- Carey, J., Bryndy, J., Wolfova, J., Grandori, R., Gustavsson, T., Ettrich, R., et al. (2007) WrbA bridges bacterial flavodoxins and eukaryotic NAD(P)H:quinone oxidoreductases. *Protein Sci.* **16**, 2301–2305
- McIver, L., LeadBeater, C., Campopiano, D. J., Baxter, R. L., Daff, S. N., Chapman, S. K., et al. (1998) Characterization of flavodoxin NADP+ oxidoreductase and flavodoxin; key components of electron transfer in *Escherichia coli*. *Eur. J. Biochem.* **257**, 577–585
- Sibille, N., Blackledge, M., Brutscher, B., Covès, J., and Bersch, B. (2005) Solution structure of the Sulfite reductase flavodoxin-like domain from *Escherichia coli*. *Biochemistry* **44**, 9086–9095
- Perez-Dorado, I., Bortolotti, A., Cortez, N., and Hermoso, J. A. (2013) Structural and phylogenetic analysis of *Rhodobacter capsulatus* NifF: uncovering general features of nitrogen-fixation (nif)-Flavodoxins. *Int. J. Mol. Sci.* **14**, 1152–1163
- Cannon, M., Cannon, F., Buchanan-Wollaston, V., Ally, D., Ally, A., and Beynon, J. (1988) The nucleotide sequence of the nifJ gene of *Klebsiella pneumoniae*. *Nucleic Acids Res.* **16**, 11379

20. Shah, V. K., Stacey, G., and Brill, W. J. (1983) Electron transport to nitrogenase: purification and characterisation of pyruvate:flavodoxin oxidoreductase, the *nifH* gene product. *J. Biol. Chem.* **258**, 12064–12068
21. Wang, Z.-Q., Lawson, R. J., Buddha, M. R., Wei, C.-C., Crane, B. R., Munro, A. W., *et al.* (2007) Bacterial flavodoxins support nitric oxide production by *Bacillus subtilis* nitric-oxide synthase. *J. Biol. Chem.* **282**, 2196–2202
22. Birch, O. M., Fuhrmann, M., and Shaw, N. M. (1995) Biotin synthase from *Escherichia coli*, an investigation of the low molecular weight and protein components required for activity *in Vitro*. *J. Biol. Chem.* **270**, 19158–19165
23. Ifuku, O., Koga, N., Haze, S., Kishimoto, J., and Wachi, Y. (1994) Flavodoxin is required for conversion of dethiobiotin to biotin in *Escherichia coli*. *Eur. J. Biochem.* **224**, 173–178
24. Hall, D. A., Van der Kooi, C. W., Stasik, C. N., Stevens, S. Y., Zuiderweg, E. R. P., and Matthews, R. G. (2001) Mapping the interactions between flavodoxin and its physiological partners flavodoxin reductase and cobalamin-dependent methionine synthase. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 9521–9526
25. Fujii, K., and Huennekens, F. M. (1974) Activation of methionine synthetase by a reduced Triphosphopyridine nucleotide-dependent flavoprotein system. *J. Biol. Chem.* **249**, 6745–6753
26. Gutekunst, K., Chen, X., Schreiber, K., Kaspar, U., Makam, S., and Appel, J. (2014) The bidirectional NiFe-hydrogenase in *synechocystis* sp PCC 6803 is reduced by flavodoxin and ferredoxin and is essential under mixotrophic, nitrate-limiting conditions. *J. Biol. Chem.* **289**, 1930–1937
27. Blamey, J. M., and Adams, M. W. W. (1994) Characterisation of an ancestral type of pyruvate ferredoxin oxidoreductase from the hyperthermophilic bacterium, *Thermotoga maritima*. *Biochemistry* **33**, 1000–1007
28. Blaschkowski, H. P., Neuer, G., Ludwig-Festl, M., and Knappe, J. (1982) Routes of flavodoxin and ferredoxin reduction in *Escherichia coli*. CoA-acylating pyruvate: flavodoxin and NADPH: flavodoxin oxidoreductases participating in the activation of pyruvate formate-lyase. *Eur. J. Biochem.* **123**, 563–569
29. Hughes, N. J., Clayton, C. L., Chalk, P. A., and Kelly, D. J. (1998) *Helicobacter pylori* porCDAB and oorDABC genes encode distinct pyruvate: flavodoxin and 2-oxoglutarate:acceptor oxidoreductases which mediate electron transport to NADP. *J. Bacteriol.* **180**, 119–1128
30. Hawkes, D. B., Adams, G. W., Burlingame, A. L., Ortiz de Montellano, P. R., and De Voss, J. J. (2002) Cytochrome P450 cin (CYP176A), isolation, expression, and characterization. *J. Biol. Chem.* **277**, 27725–27732
31. Bruender, N. A., Young, A. P., and Bandarian, V. (2015) Chemical and biological reduction of the radical SAM enzyme 7-Carboxy-7-deaza-guanine synthase. *Biochemistry* **54**, 2903–2910
32. Tejero, J., Hannibal, L., Mustovich, A., and Stuehr, D. J. (2010) Surface charges and regulation of FMN to heme electron transfer in nitric-oxide synthase. *J. Biol. Chem.* **285**, 27232–27240
33. Aoki, M., Ishimori, K., and Morishima, I. (1998) Roles of negatively charged surface residues of putidaredoxin in interactions with redox partners in P450cam monooxygenase system. *Biochim. Biophys. Acta* **1386**, 157–167
34. Wan, J. T., and Jarrett, J. T. (2002) Electron acceptor specificity of ferredoxin (flavodoxin): NADP⁺ oxidoreductase from *Escherichia coli*. *Arch. Biochem. Biophys.* **406**, 116–126
35. Smillie, R. M. (1965) Isolation of phytoflavin, a flavoprotein with chloroplast ferredoxin activity. *Plant Physiol.* **40**, 1124–1128
36. López-Llano, J., Maldonado, S., Bueno, M., Lostao, A., Ángeles-Jiménez, M., Lillo, M. P., *et al.* (2004) The long and short flavodoxins I. The role of the differentiating loop in apoflavodoxin structure and FMN binding. *J. Biol. Chem.* **279**, 47177–47183
37. Cao, P., Cao, D. F., Si, L., Su, X. D., Chang, W. R., Liu, Z. F., *et al.* (2020) Structural basis for energy and electron transfer of the photosystem I-IsiA-flavodoxin supercomplex. *Nat. Plants* **6**, 167–176
38. Pechter, K. B., Gallagher, L., Pyles, H., Manoil, C. S., and Harwood, C. S. (2015) Essential genome of the metabolically versatile alphaproteobacterium *Rhodospseudomonas palustris*. *J. Bacteriol.* **198**, 867–876
39. Ye, Q., Hu, Y., and Jin, C. (2014) Conformational dynamics of *Escherichia coli* flavodoxins in apo- and holo-states by solution NMR spectroscopy. *PLoS One* **9**, e103936
40. Varner, T. A., Mohamed-Raseek, N., and Miller, A.-F. (2021) Assignments of ¹⁹F NMR resonances and exploration of dynamics in a long-chain flavodoxin. *Achiv. Biochem. Biophys.* **703**, 108839
41. Banci, L., Bertini, I., Calderone, V., Ciofi-Baffoni, S., Giachetti, A., Jaiswal, D., *et al.* (2013) Molecular view of an electron transfer process essential for iron–sulfur protein biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7136–7141
42. Segal, H. M., Spatzal, T., Hill, M. G., Udit, A. K., and Rees, D. C. (2017) Electrochemical and structural characterization of *Azotobacter vinelandii* flavodoxin II. *Protein Sci.* **26**, 1984–1993
43. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797
44. Romao, C. V., Vicente, J. B., Borges, P. T., Victor, B. L., Lamosa, P., Silva, E., *et al.* (2016) The crystallographic structure of Flavobredoxin from *Escherichia coli*. *J. Mol. Biol.* **428**, 4686–4707
45. Silaghi-Dumitrescu, R., Kurtz, D. M., Jr., Ljungdahl, L. G., and Lanzetta, W. N. (2005) X-ray crystal structures of Moorella thermoacetica FprA. Novel diiron site structure and mechanistic insights into a scavenging nitric oxide reductase. *Biochemistry* **44**, 6492–6501
46. Hickey, D. P., Lim, K., Cai, R., Patterson, A. R., Yuan, M., Sahin, S., *et al.* (2018) Pyrene hydrogel for promoting direct bio-electrochemistry: ATP-independent electroenzymatic reduction of N₂. *Chem. Sci.* **9**, 5172–5177
47. Imam, H. T., Marr, P. C., and Marr, A. C. (2021) Enzyme entrapment, biocatalyst immobilization without covalent attachment. *Green Chem.* **23**, 4980–5005
48. Giroud, F., and Minter, S. D. (2013) Anthracene-modified pyrenes immobilized on carbon nanotubes for direct electroreduction of O₂ by laccase. *Electrochem. Commun.* **34**, 157–160
49. Hoover, D. M., Drennan, C. L., Metzger, A. L., Osborne, C., Weber, C. H., Patridge, K. A., *et al.* (1999) Comparisons of wild-type and mutant flavodoxins from *Anacystis nidulans*. Structural determinants of the redox potentials. *J. Mol. Biol.* **294**, 725–743
50. Lostao, A., Gomez-Moreno, C., Mayhew, S. G., and Sancho, J. (1997) Differential stabilization of the three FMN redox forms by tyrosine 94 and tryptophan 57 in flavodoxin from *Anabaena* and its influence on the redox potentials. *Biochemistry* **36**, 14334–14344
51. Zhang, Y. (2017) *Methylophilus in a Groundwater-Derived Drinking Water System*, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL
52. Seyedasaymdest, M. R., Reece, S. Y., Nocera, D. G., and Stubbe, J. (2006) Mono-, di-, tri and tetra-substituted fluorotyrosines: new probes for enzymes that use tyrosyl radicals in catalysis. *J. Am. Chem. Soc.* **128**, 1569–1579
53. Sitbon, E., and Pietrovski, S. (2007) Occurrence of protein structure elements in conserved sequence regions. *BMC Struct. Biol.* **7**, 3
54. Kozłowski, L. P. (2017) Proteome-pI: proteome isoelectric point database. *Nucleic Acids Res.* **45**, D1112–D1116
55. Shingate, P., and Sowdhamini, R. (2012) Analysis of domain-swapped oligomers reveals local sequence preferences and structural imprints at the linker regions and swapped interfaces. *PLoS one* **7**, e39305
56. Pueyo, J. J., Gómez-Moreno, C., and Mayhew, S. G. (1991) Oxidation-reduction potentials of ferredoxin-NADP⁺ reductase and flavodoxin from *Anabaena* PCC 7119 and their electrostatic complexes. *Eur. J. Biochem.* **202**, 1065–1071
57. Saen-oon, S., de Vaca, I. C., Masone, D., Medina, M., and Guallar, V. (2015) A theoretical multiscale treatment of protein–protein electron transfer: the ferredoxin/ferredoxin-NADP⁺ reductase and flavodoxin/ferredoxin-NADP⁺ reductase systems. *Biochim. Biophys. Acta* **1847**, 1530–1538
58. Kurisu, G., Kusunoki, M., Katoh, E., Yamazaki, T., Teshima, K., Onda, Y., *et al.* (2001) Structure of the electron transfer complex between ferredoxin and ferredoxin-NADP⁺ reductase. *Nat. Struct. Biol.* **8**, 117–121
59. Ubbink, M. (2012) Dynamics in transient complexes of redox proteins. *Biochem. Soc. Trans.* **40**, 415–418

60. Ludwig, M. L., Patridge, K. A., Metzger, A. L., Dixon, M. M., Eren, M., Feng, Y., *et al.* (1997) Control of oxidation-reduction potentials in flavodoxin from *Clostridium beijerinckii*: the role of conformational changes. *Biochemistry* **36**, 1259–1280
61. Liu, S. J., Chen, Y., Du, T., Zhao, W., Liu, X., Zhang, H., *et al.* (2023) A dimer-monomer transition captured by the crystal structures of cyanobacterial apo flavodoxin. *Biochem. Biophys. Res. Commun.* **639**, 134–141
62. Hrovat, A., Blümel, M., Löhr, F., Mayhew, S. G., and Rüterjans, H. (1997) Backbone dynamics of oxidized and reduced *D. vulgaris* flavodoxin in solution. *J. Biomol. NMR* **10**, 53–62
63. Mutoh, R., Muraki, N., Shinmura, K., Kubota-Kawai, H., Lee, Y.-H., Nowaczyk, M. M., *et al.* (2015) X-ray structure and nuclear magnetic resonance analysis of the interaction sites of the Ga-substituted cyanobacterial ferredoxin. *Biochemistry* **54**, 6052–6061
64. Crain, A. V., and Broderick, J. B. (2014) Flavodoxin cofactor binding induces structural changes that are required for protein-protein interactions with NADP⁺ oxidoreductase and pyruvate formate-lyase activating enzyme. *Biochim. Biophys. Acta* **1834**, 2512–2519
65. Porter, T. D., and Kasper, C. B. (1986) NADPH-cytochrome P-450 oxidoreductase: flavin mononucleotide and flavin adenine dinucleotide domains evolved from different flavoproteins. *Biochemistry* **25**, 1682–1687
66. Ingelman, M., Bianchi, V., and Eklund, H. (1997) The three-dimensional structure of flavodoxin reductase from *Escherichia coli* at 1.7 Å resolution. *J. Mol. Biol.* **268**, 147–157
67. Xia, C., Rwere, R., Im, S., Shen, A. L., Waskell, L., and Kim, J. J. (2018) Structural and kinetic studies of Asp632 mutants and fully- reduced NADPH-cytochrome P450 oxidoreductase define the role of Asp632 loop dynamics in control of NADPH binding and hydride transfer. *Biochemistry* **57**, 945–962
68. Urban, P., and Pompon, D. (2022) Confrontation of AlphaFold models with experimental structures enlightens conformational dynamics supporting CYP102A1 functions. *Sci. Rep.* **12**, 15982
69. Govindaraj, S., and Poulos, T. L. (1997) The domain architecture of cytochrome P450bm-3. *J. Biol. Chem.* **272**, 7915–7921
70. Neeli, R., Girvan, H. M., Lawrence, A., Warren, M. J., Leys, D., Scrutton, N. S., *et al.* (2005) The dimeric form of flavocytochrome P450 BM3 is catalytically functional as a fatty acid hydroxylase. *FEBS Lett.* **579**, 5582–5588
71. Ermakova, M., Battchikova, N., Richaud, P., Leino, H., Kosourov, S., Isojäävi, J., *et al.* (2014) Heterocyst-specific flavodiiron protein Flv3B enables oxic diazotrophic growth of the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 11205–11210
72. Silaghi-Dumitrescu, R., Coulter, E. D., Das, A., Ljungdahl, L., Jameson, G. N. L., Huynh, B. H., *et al.* (2003) A flavodi-iron protein and high molecular weight rubredoxin from *Moorella thermoacetica* with nitric oxide reductase activity. *Biochemistry* **42**, 2806–2815
73. Warren, J. J., Ener, M. E., Vlcek, A., Jr., Winkler, J. R., and Gray, H. B. (2012) Electron hopping through proteins. *Coord. Chem. Rev.* **256**, 2478–2487
74. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol.* **177**, 44–73
75. Antharavally, B. S., Mallia, K. A., Rangaraj, P., Haney, P., and Bell, P. A. (2009) Quantitations of proteins using a dye-metal-based colorimetric protein assay. *Anal. Biochem.* **385**, 342–345
76. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214
77. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
78. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
79. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
80. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., *et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367
81. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
82. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21
83. Bliven, S., Lafita, A., Parker, A., Capitani, G., and Duarte, J. M. (2018) Automated evaluation of quaternary structures from protein crystals. *PLoS Comput. Biol.* **14**, e1006104
84. Zallot, R., Oberg, N., and Gerlt, J. A. (2019) The EFI web resource for genomic enzymology tools: leveraging protein, genome, and meta-genome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **58**, 4169–4182
85. Oberg, N., Zallot, R., and Gerlt, J. A. (2023) EFI-EST, EFI-GNT, and EFI-CGFP: enzyme function initiative (EFI) web Resource for genomic enzymology tools. *J. Mol. Biol.* **435**, 168018
86. van Mierlo, C. P. M., Lijnzaad, P., Vervoort, J., Mueller, F., Berendsen, H. J., and de Vlieg, J. (1990) Tertiary structure of two-electron reduced Megasphaera elsdenii flavodoxin and some implications, as determined by two-dimensional 1H NMR and restrained molecular dynamics. *Eur. J. Biochem.* **194**, 185–198
87. Thomsen, M. C. F., and Nielsen, M. (2012) Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Res.* **40**, W281–W287
88. Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C., and Ferrin, T. E. (2006) Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **7**, 339
89. Sevrioukova, I. F., Li, H., Zhang, H., Peterson, J. A., and Poulos, T. L. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 1863–1868
90. Alagaratnam, S., van Pouderoyen, G., Pijning, T., Dijkstr, B. W., Cavazzini, D., Rossi, G. L., *et al.* (2005) A crystallographic study of Cys69Ala flavodoxin II from *Azotobacter vinelandii*: structural determinants of redox potential. *Prot. Sci.* **14**, 2284–2295
91. Steensma, E., Heering, H. A., Hagen, W. R., and Van Mierlo, C. P. (1996) Redox properties of wild-type, Cys69Ala, and Cys69Ser *Azotobacter vinelandii* flavodoxin II as measured by cyclic voltammetry and EPR spectroscopy. *Eur. J. Biochem.* **235**, 167–172
92. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., *et al.* (2014) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303
93. Silaghi-Dumitrescu, R., Kurtz, D. M., Jr., Ljungdahl, L. G., and Lanzilotta, W. N. (2005) X-ray crystal structures of *Moorella thermoacetica* FprA. Novel diiron site structure and mechanistic insights into a scavenging nitric oxide reductase. *Biochemistry* **44**, 6492–6501
94. Prakash, D., Iyer, P. R., Suharti, S., Walters, K. A., Santiago-Martinez, G. M., Golbeck, J. H., *et al.* (2019) Structure and function of an unusual flavodoxin from the domain *Archaea*. *Proc. Nat. Acad. Sci. U. S. A.* **116**, 25917–25922
95. Hsieh, Y.-C., Chia, T. S., Fun, H.-K., and Chen, C.-J. (2013) Crystal structure of dimeric flavodoxin from *Desulfovibrio gigas* suggests a potential binding region for the electron-transferring partner. *Int. J. Mol. Sci.* **14**, 1667–1683
96. Schopfer, L. M., Ludwig, M. L., and Massey, V. (1991) A working proposal for the role of the apoprotein in determining the redox potential of the flavin in flavoproteins: correlations between potentials and flavin

- pKs. In: Curti, B., Ronchi, S., Zanetti, G., eds. *Flavins and Flavoproteins* 1990, Walter de Gruyter, Berlin, Germany: 399–404
97. Mayhew, S. G. (1999) The effects of pH and semiquinone formation on the oxidation- reduction potentials of flavin mononucleotide: a reappraisal. *Eur. J. Biochem.* **265**, 698–702
 98. Silaghi-Dumitrescu, R., Coulter, E. D., Das, A., Ljungdahl, L., Jameson, G. N. L., Huynh, B. H., *et al.* (2003) A flavodi-iron protein and high molecular weight rubredoxin from *Moorella thermoacetica* with nitric oxide reductase activity. *Biochemistry* **42**, 2806–2815
 99. Folgosa, F., Martins, M. C., and Teixeira, M. (2018) Diversity and complexity of flavodiiron NO/O₂ reductases. *FEMS Microbiol. Lett.* **365**, fnx267
 100. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., *et al.* (2004) UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612
 101. Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S., and Kim, J. J. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 8411–8416