

SOD1 Mutations Targeting Surface Hydrogen Bonds Promote Amyotrophic Lateral Sclerosis without Reducing Apo-state Stability^{*[5]}

Received for publication, November 18, 2009, and in revised form, February 5, 2010. Published, JBC Papers in Press, February 26, 2010, DOI 10.1074/jbc.M109.086074

Roberth Byström[‡], Peter M. Andersen[§], Gerhard Gröbner[‡], and Mikael Oliveberg^{¶1}

From the [¶]Department of Biochemistry and Biophysics, Arrhenius Laboratories of Natural Sciences, Stockholm University, S-106 91 Stockholm, Sweden and the Departments of [‡]Chemistry and [§]Clinical Neuroscience, Umeå University, S-901 87 Umeå, Sweden

In good accord with the protein aggregation hypothesis for neurodegenerative diseases, ALS-associated SOD1 mutations are found to reduce structural stability or net repulsive charge. Moreover there are weak indications that the ALS disease progression rate is correlated with the degree of mutational impact on the apoSOD1 structure. A bottleneck for obtaining more conclusive information about these structure-disease relationships, however, is the large intrinsic variability in patient survival times and insufficient disease statistics for the majority of ALS-provoking mutations. As an alternative test of the structure-disease relationship we focus here on the SOD1 mutations that appear to be outliers in the data set. The results identify several ALS-provoking mutations whose only effect on apoSOD1 is the elimination or introduction of a single charge, *i.e.* D76V/Y, D101N, and N139D/K. The thermodynamic stability and folding behavior of these mutants are indistinguishable from the wild-type control. Moreover, D101N is an outlier in the plot of stability loss *versus* patient survival time by having rapid disease progression. Common to the identified mutations is that they truncate conserved salt-links and/or H-bond networks in the functional loops IV or VII. The results show that the local impact of ALS-associated mutations on the SOD1 molecule can sometimes overrun their global effects on apo-state stability and net repulsive charge, and point at the analysis of property outliers as an efficient strategy for mapping out new ALS-provoking features.

As in other soluble proteins, the charged and polar side chains of SOD1 are found in connection to the protein surface where they protrude freely into the solvent or are involved in solvent-accessible salt links and hydrogen bonds (Fig. 1). Replacement of such surface groups by point mutation has typically small effects on protein folding and stability. Even so, charged side chains have been observed to play key roles in controlling protein-protein interactions and aggregation. This control can either be exerted in the con-

text of the folded structure by, *e.g.* providing edge protection of β -sheets (1, 2), or by modulation the aggregation propensity of the unfolded chain (1, 3). In this perspective it is interesting that 48 out of the 142 SOD1 mutations that have so far been linked to the neurodegenerative disease amyotrophic lateral sclerosis (ALS)² involve charge replacements (4) (supplemental Table S1). Like most proteins, SOD1 carries a net negative charge under physiological conditions, which assure some degree of electrostatic repulsion between the molecular components in the crowded cellular interior. Strikingly, 83% of the ALS-mutations targeting SOD1 charges are found to decrease this net repulsive charge (5) and genetic data suggest the loss of a single negative charge is enough to trigger ALS (5). At the other end of the spectrum are the ALS-associated mutations of hydrophobic residues in the SOD1 interior and of the loop glycines (Fig. 1). The common effect of these mutations is to destabilise the SOD1 framework by shifting the folding equilibrium toward the denatured state. For some mutants, *e.g.* A4V and G93A, the protein is even destabilised to the extent that the nascent chain cannot fold under physiological conditions without oxidation of the Cys⁵⁷–Cys¹⁴⁶ disulfide bridge or acquisition of the native metals (6). Taken together, this tendency of the ALS-associated SOD1 mutations to reduce either repulsive charge or structural stability point at protein aggregation as a deterministic factor in ALS (5, 7). This physical-chemical link to protein aggregation complies also with the occurrence of pathological SOD1 inclusions in the spinal cord of ALS patients (8–11). Another conspicuous feature of the ALS mechanism is that several of the SOD1 mutations show a characteristic uniform disease progression. For example, the survival time after symptomatic onset for patients carrying the A4V is invariably short (<2 years), whereas the survival time for patients with the H46R is always very long (>>10 years). On this basis several investigations have been directed to examine if there is a coupling between the physical-chemical properties of the ALS-provoking SOD1 mutations and the severity of disease (12–14). That is, are the most perturbing mutations also the ones that cause the worse damage? Regarding the age of onset there is no apparent trend, most mutations for which reliable data are avail-

* This work was supported in part by the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Bertil Hällsten Foundation, and Hjärnfonden.

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Tables S1 and S2.

¹ To whom correspondence should be addressed. Tel: 46-8-162459; Fax: 46-8-153679; E-mail: mikael.oliveberg@dbb.su.se.

² The abbreviations used are: ALS, amyotrophic lateral sclerosis; MES, 4-morpholineethanesulfonic acid; PDB, Protein Data Bank.

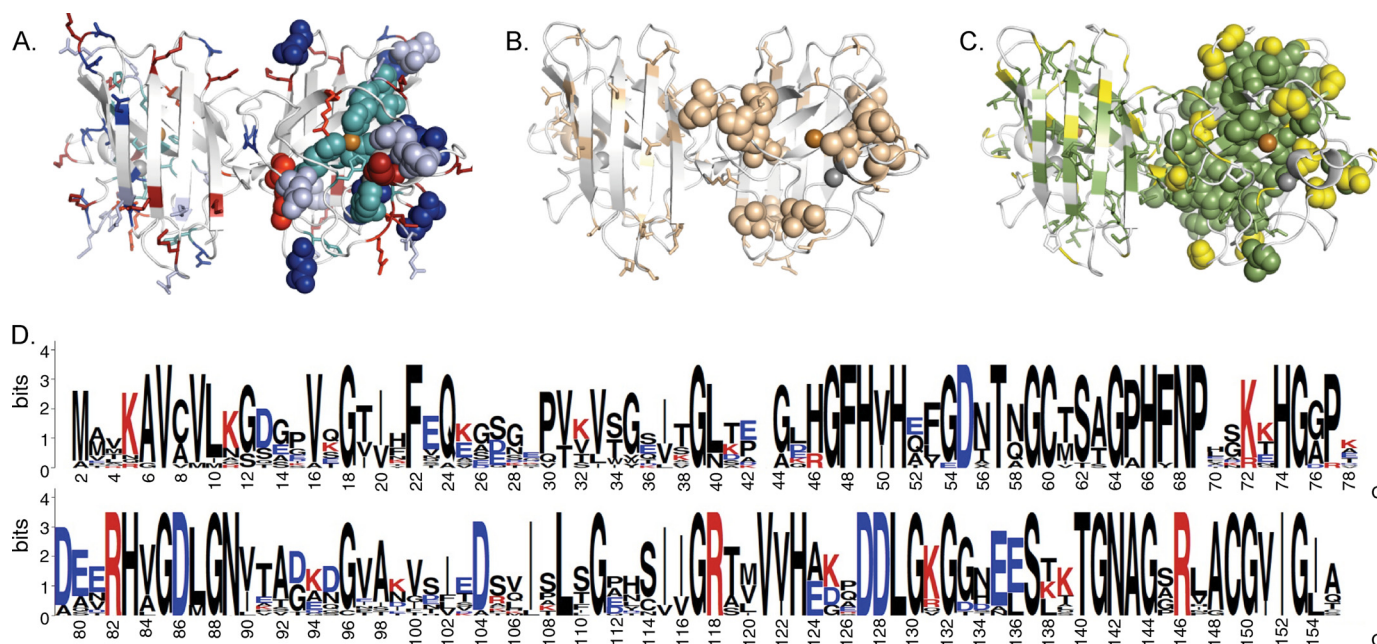


FIGURE 1. The distribution of charged, polar, and hydrophobic amino acids in the SOD1 structure (PDB entry 1HL5) is shown as sticks. Positions with ALS-associated mutations are marked as spheres in the right monomer only. *A*, SOD1 dimer showing the positions of charged side chains, D (blue), E (light blue), K (red), and R (bright red), and histidines (cyan). *B*, structural positions of the polar (beige) side chains (S, T, C, N, and Q). *C*, structural positions of the hydrophobic (green) side chains (A, V, F, M, I, W, and L). Proline (white) and glycine (yellow). *D*, SOD1 amino acid sequence indicating the degree of side chain conservation across 17 eukaryotic species as visualized by Homologene (53). The degree of conservation is measured in bits ($\log_2(20) = 4.3$ bits, where 20 is the number of possible amino acids in a peptide chain). For each sequence position there is a stack of amino acid occurrences. The height of each stack indicates the sequence conservation for that position (54). The height of each letter reflects the relative frequency for the amino acid in question and the letter at the top of each stack is the most frequent in that position. Multiple alignments were made with the T-coffee software using default parameters (55). The consensus sequence was made with the WebLogo software (54).

able has a mean age of onset of 46–47 years (14), indicating the ALS mechanism in this respect is different from that of Lund-Huntington disease where the length of the poly(Q) additions correlates with age of onset of clinical symptoms (15, 16). When it comes to the disease progression, however, the thermodynamic stability of the apoSOD1 species hints at a weak correlation with survival time after first diagnosis for a subset of non-charged ALS mutations selected on the basis of good clinical statistics (17). This trend has also been discerned with a more extensive set of ALS mutations even though the scatter of the data is close to abbreviate the significance of the signal (13). The main problem is uncertain survival statistics because of low patient numbers. For the majority (76%) of the ALS-associated SOD1 mutations there are less than 5 reported disease cases and this shortage of data is particularly alarming for mutations with long survival times (supplemental Table S1). In addition to this statistical uncertainty, however, there seems also to be mutations that simply break the pattern. A good example of such a statistically safe outlier is the mutation D101N. This mutation yields a thermal melting transition close to that of the wild-type protein and is still associated with patient survival times of <5 years ($n = 17$). In sharp contrast, the more destabilizing ALS mutation D90A displays typically survival times >5 years ($n = 15$) despite having the same effect on the net repulsive charge. Both mutations reduce the charge of the apoSOD1 monomer from -6 to -5 accordingly, it is reasonable to assume that these differences in survival time stem from specific effects on the SOD1 molecule or its interaction with cellular components that go beyond any underlying

dependence on the global parameters net charge and thermodynamic stability. In this study, we have systematically tried to verify and extend the set of property outliers to see if these deviating SOD1 mutations can reveal new clues to the ALS mechanism. The analysis identifies 5 ALS-associated charge mutations with apstate stabilities and global folding transitions identical to those of the wild-type protein, *i.e.* D76V, D76Y, D101N, N139D, and N139K. The results provide direct experimental support for the genetic indication that reduction of the net charge by just one unit induces ALS with high penetrance. Moreover, the data point at an additional molecular factor in the mechanism of SOD-induced ALS: the rupture of conserved hydrogen-bond networks at the SOD1 surface.

EXPERIMENTAL PROCEDURES

All experiments were done on apo-protein under oxidizing conditions with the intramolecular disulfide bridge between Cys⁵⁷ and Cys¹⁴⁶ kept intact.

Protein Preparation—All mutants were made on a background of C6A/C111A and for the monomers on C6A/C111A/F50E/G51E. The proteins were coexpressed with the copper chaperone (yCCS) as described in Ref. 12. The apo-protein was prepared as described in Ref. 6. The standard buffer was 10 mM MES (pH 6.3) with 10 mM EDTA to maintain the proteins as metal-free.

Kinetic Measurements—Stopped-flow measurements were performed on a SX-17MV stopped-flow spectrofluorometer (Applied Photophysics, Leatherhead, UK) with excitation at 280 nm and emission detection above 320 nm using a cut-off filter. For slow relaxations, *i.e.* $k < 0.003$ s⁻¹, additional mea-

ALS Mutations Targeting SOD1 Hydrogen Bonds

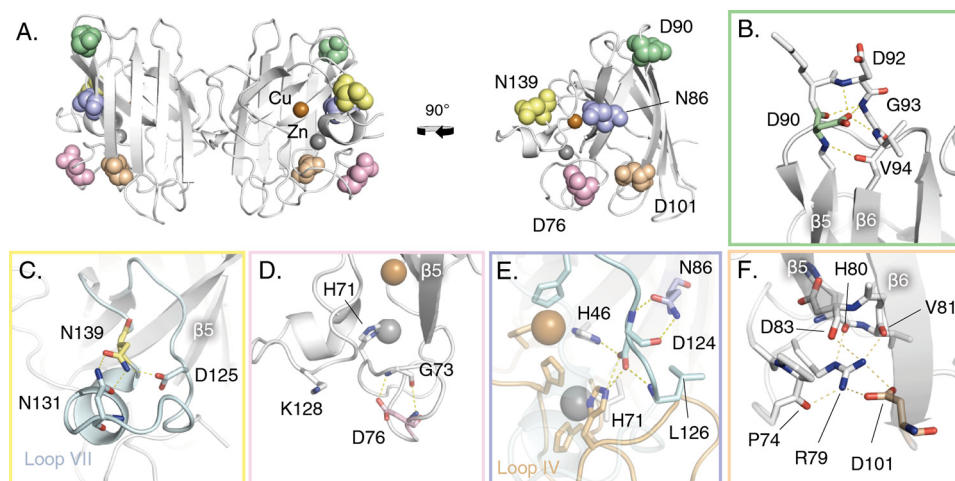


FIGURE 2. **Structural positions of the ALS-associated SOD1 mutations examined in this study.** Each mutant is color coded with a corresponding color-coded close-up on the affected hydrogen bonds. A, front view of SOD1 dimer and side view of monomer (PDB entry 1HL5). B–F, close-ups of the local structural environments of Asp⁹⁰, Asn¹³⁹, Asp⁷⁶, Asn⁸⁶, and Asp¹⁰¹, respectively.

measurements were done by manual mixing with detection on a FP-6500 spectrofluorometer (Jasco). All measurements were done at 25 °C and at a final SOD1 concentration of 4 μ M monomer. Urea was used as denaturant (ultra PURE; MP Biomedicals, Solon, OH).

Data Analysis—Under the assumption of two-state kinetics for the apoSOD1 monomer (6), protein stability was derived from chevron data according to Equation 1,

$$\Delta G_{D-M} = -RT \ln K_{D-M} = -RT \ln (k_u^{H_2O} / k_f^{H_2O}) \quad (\text{Eq. 1})$$

where ΔG_{D-M} is the protein stability, $K_{D-M} = [D]/[M]$, k_u is the unfolding rate constant, and k_f is the refolding rate constant. The plots of $\log k_f$ and $\log k_u$ versus [urea] were fitted by Refs. 6 and 18) in Equation 2,

$$\log k_{\text{obs}} = \log(k_f + k_u) = \log(10^{\log k_f^{H_2O} + m_f[\text{urea}]} + 10^{\log k_u^{H_2O} + m_u[\text{urea}]}) \quad (\text{Eq. 2})$$

where $k_f^{H_2O}$ and $k_u^{H_2O}$ are the rate constants at [urea] = 0 M, and m_f and m_u are the slopes of the refolding and unfolding limbs, respectively. The stability changes caused by point mutations were calculated from Equation 3,

$$\Delta \Delta G_{D-M}^{\text{mut}} = \Delta G_{D-M}^{\text{pWT}} - \Delta G_{D-M}^{\text{mut}} = -2.3RT (\Delta \log k_f^{H_2O} - \Delta \log k_u^{H_2O}) \quad (\text{Eq. 3})$$

where $\Delta G_{D-M}^{\text{pWT}}$ is the stability of the pseudo-WT and $\Delta G_{D-M}^{\text{mut}}$ is the stability of the mutant protein. $\Delta \log k_f^{H_2O}$ and $\Delta \log k_u^{H_2O}$ are the mutant-induced changes of the rate constants. The changes in stability upon mutation of dimeric apoSOD was calculated from 6 in Equation 4,

$$\Delta \Delta G_{2D-M}^{\text{mut}} = \Delta \Delta G_{D-M}^{\text{mut}} + (-2.3RT \cdot (\log k_d^{\text{pWT}} - \log k_d^{\text{mut}})) \quad (\text{Eq. 4})$$

where $\log k_d$ is the dimer dissociation rate constant at 5.8 M [urea]. For data analysis, we used KALEIDAGRAPH software (Synergy Software, Reading, PA).

Normalization of Protein Stabilities—To be able to compare mutant stabilities measured by denaturant-induced and thermal unfolding we normalized the data as follows. For denaturant-induced measurements, the normalized destabilization upon point mutation ($\Delta \Delta G^{\text{norm}}$) was estimated as Equation 5,

$$\Delta \Delta G^{\text{norm}} = 1 - (\Delta \Delta G^{\text{obs}} - \Delta \Delta G^{\text{MIN}}) / (\Delta \Delta G^{\text{MAX}} - \Delta \Delta G^{\text{MIN}}), \quad (\text{Eq. 5})$$

where $\Delta \Delta G^{\text{obs}}$ is the observed destabilization for the mutation in question, and $\Delta \Delta G^{\text{MAX}}$ and $\Delta \Delta G^{\text{MIN}}$ are the maximum and minimum values of the dataset taken from Table 2 in Ref. 17. For the thermal data, $\Delta \Delta G^{\text{norm}}$ was correspondingly derived as Equation 6,

$$\Delta \Delta G^{\text{norm}} = 1 - (\Delta T_m^{\text{obs}} - \Delta T_m^{\text{MIN}}) / (\Delta T_m^{\text{MAX}} - \Delta T_m^{\text{MIN}}), \quad (\text{Eq. 6})$$

where the changes in melting temperatures upon point mutation ΔT_m^{obs} , ΔT_m^{MAX} , and ΔT_m^{MIN} were averages of the apo-monomer data in Refs. 19–23.

RESULTS AND DISCUSSION

Screening of ALS Mutations at the SOD1 Surface—To identify ALS-associated SOD1 mutations with minimal impact on structural stability we focused on a set of surface-associated positions at maximum distance from the dimer interface, *i.e.* D76V, D76Y, N86D, N86K, N86S, D90V, D90A, D101N, D101G, N139D, and N139K (Fig. 2). The apparent role of these residues is to form weak, peripheral solvent-accessible hydrogen bonds or salt links, and their mutation seems overall to cause minimal steric conflicts. The effect of the mutations on the structural stability and folding of the apoSOD1 molecule was determined by chevron analysis (6, 18). A particular advantage of this approach is that stability changes and perturbations of the folding process can be observed simultaneously from changes of folding and unfolding rate constants. Moreover, the use of urea as a denaturant reduces the risk of irreversibility artifacts because of aggregation of the unfolded protein. To further reduce the risk of erroneous disulfide cross-linking all experiments were done on the background C6A/C111A. Although these substitutions are somewhat destabilizing to the SOD1 structure this contribution cancels upon comparison of the pseudo-wildtype and mutant proteins (6). Finally, to assure monomeric apo-states, we added to all constructs the dimer-splitting substitutions F50E/G51E (24). Again, the small absolute effects of these substitutions are additive and cancel effectively upon comparative analysis of the pseudo-wild-type and mutant proteins (6).

Folding of apoSOD^{pWT}—The folding of wild-type apoSOD1, as well as the C6A/C111A pseudo wild-type variant (apoSOD^{pWT}),

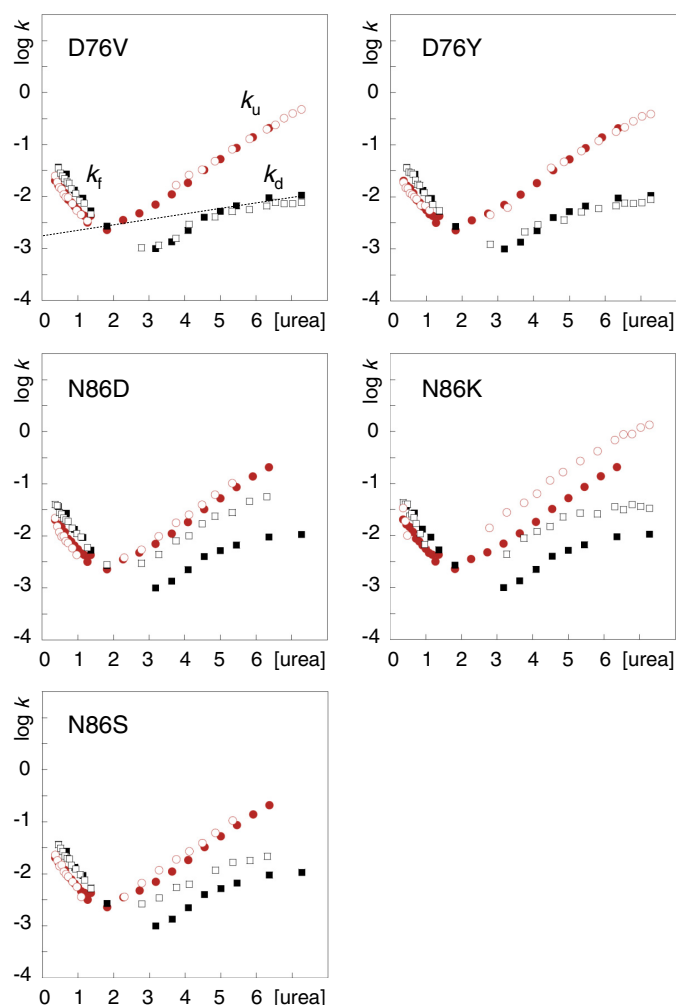
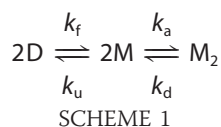


FIGURE 3. **Chevron plots of D76V, D76Y, N86D, N86K, and N86S.** k_f and k_u are the rate constants for folding and refolding, respectively, and the *dashed line* shows the urea dependence of $\log k_d$. apoSOD1^{WT}_{mono} (●), apoSOD1^{mut}_{mono} (○), apoSOD1^{WT}_{dimer} (■) and apoSOD1^{mut}_{dimer} (□). The units of [urea] and k are M and s^{-1} , respectively.

follow a three-state process (25), in which the monomer forms independently before the assembly of the final homodimeric structure (6, 26) in Scheme 1,



where D is the denatured state, M is the folded monomer, M_2 is the folded homodimer, k_f and k_u are the refolding and unfolding rate constants, and k_a and k_d the rate constants for dimer association and dissociation, respectively. Like other small globular proteins, the apoSOD1 monomer folds in a cooperative, two-state manner characterized by a classical V-shaped chevron plot (Fig. 3), where the ratio of k_u and k_f measures monomer stability (Equation 1). This cooperative formation of the apoSOD1 monomer is also the rate-limiting step for dimer formation. The small off-set of the refolding limbs ($\log k_f$ versus [urea]) between the monomeric and dimeric SOD1 variants is caused by the dimer-splitting substitutions F50E/G51E (6).

With respect to unfolding reaction, the apoSOD1 dimers show a downward kink in the chevron plots at 5 M urea where dimer dissociation (k_d) becomes the rate-limiting step. Accordingly, any mutational perturbation of the stability or folding of the SOD1 monomers and dimers can be identified by characteristic alterations of their chevron plots (17).

D76V and D76Y—Our results show that the chevron plots of both the monomeric and homodimeric forms of apoSOD1^{D76V} and apoSOD1^{D76Y} are identical to those of the pseudo wild-type protein (Fig. 3). This suggests that mutational substitution of the Asp⁷⁶ side chain at the surface of loop IV is structurally benign and has no significant impact on SOD1 stability of the way the protein attains its folded state *in vitro*. Yet, Asp⁷⁶ displays a high degree of conservation across the SOD1 family (Fig. 1) indicating that the carboxylate moiety of this side chain plays an important functional or structural role.

The most obvious structural clue to what this role could be is a hydrogen bond to Gly⁷³, close to the Zn ligand H71 (Fig. 2D). In addition, there are indications of a conserved salt link between Asp⁷⁶ and Lys¹²⁸ in loop VI (27) although this bond is missing or disorganized in several crystal structures, *e.g.* 1HL5 (28) and 1PUO (29). By these surface interactions Asp⁷⁶ could provide specificity to the geometry of the peripheral turn of loop IV, affecting for example the accessibility of proteases to the putative nick sites at Arg⁶⁹/Lys⁷⁰ and the neighboring Lys⁷⁵ (30). Even if the hydrogen bond and putative salt link do not contribute much to the thermodynamic stability of SOD1 structure they can still restrict the loop dynamics; because hydrogen bonds are inclined to be either fully formed or open to solvent interactions they provide some degree of click fit. Consistent with such an increased sensitivity to degradation, D76Y and the nearby mutation G72S show only half of the expected enzymatic turnover in erythrocytes from ALS patients (31, 32). It is also conceivable that Asp⁷⁶ can modulate macromolecular association (1, 2) or be a key contributor to the electrostatic potential around the active site.

In line with the negligible effect of D76V and D76Y on SOD1 folding and stability, the few ALS patients diagnosed with these mutations display overall very slow disease progression. The survival times after first diagnosis are 19 years ($n = 4$) and 17/5 years ($n = 2$) for D76V and D76Y, respectively. It is also notable that the disease penetrance of D76Y is incomplete (31), indicating that the mutation just reaches the pathogenic threshold. The ALS-inducing property seems in this case to be the sole loss of a negative charge.

N86D, N86K, and N86S—The three ALS-associated mutations of Asn⁸⁶ lead to characteristic perturbations the SOD1 stability that depends on the identity of the substitution (Fig. 3). N86K decreases the stability of the apoSOD1 monomer by a selective increase of the unfolding rate constant, k_u . According to standard ϕ -value analysis (18) this suggests that the mutational perturbation sets in late in the folding process and affects mainly the fully structured apo-monomer (17). The most likely cause of the perturbation is steric clashes because the introduced side chain of Lys is larger than the original Asn. Consistently, mutation of Asn⁸⁶ to the smaller side chains Asp and Ser yields no appreciable effect on the monomer stability. Even so, and more puzzlingly, all three mutations seem to weaken the

ALS Mutations Targeting SOD1 Hydrogen Bonds

strength of the dimer interface despite being located at the opposite ends the constituent monomers (Fig. 2E). The effect is manifested in an accelerated rate constant for dimer dissociation (k_d) that is similar for all three mutations (Fig. 3). An explanation for this long-range coupling to the dimer interface is hinted by two hydrogen-bond links between Asn⁸⁶ and the backbone of Asp¹²⁴ (loop VII) that, in turn, propagates to the metal ions via hydrogen bonds to the Cu ligand His⁴⁶ (β_4) and the Zn ligand His⁷¹ (loop IV) (Fig. 2E). The final coupling to the dimer interface could then be mediated by altered dynamics of loop VII to the interface strand β_8 , or via loop IV and the metal coordination. It has been demonstrated in several earlier studies that loop IV forms an integral part of the dimer interface that is modulated by both the Cys⁵⁷–Cys¹⁴⁶ disulfide cross-link and metal binding; disulfide-reduced apoSOD1 molecules lose their ability to dimerize at μM concentration (6, 33, 34). Moreover, the dimer interface and metal-binding loops are found to be coupled dynamically through a high connectivity, interaction network (35). The importance of solvent-accessible hydrogen bonds for folding and stability of loops has also been demonstrated by the insertion of artificial side chain moieties in the WW domain (36).

In terms of disease progression, however, N86D, N86S, and N86K seem to follow a simplistic trend based on the extent of mutational perturbation, albeit that the patient statistics are still too sparse to be conclusive. The single ALS patient identified with N86D, involving a putatively protective increase of the net charge is still alive four years after onset and displays slow progression. The similarly destabilized N86S, which does not alter the charge, shows a survival time after first diagnosis of 6.8 years ($n = 4$), and the most perturbing N86K, which also involves a reduction of the net negative charge, shows a survival time of 1.7 years ($n = 7$) and complete disease penetrance. Another notable feature is that N86S and N86D have incomplete disease penetrance (14),³ indicating that marginally perturbing SOD1 mutations are indeed on the verge to provoke ALS.

D90V and D90A—From the chevron data it is apparent that the two mutations in position D90 destabilize selectively the apoSOD1 monomer, whereas the dimer interface is left relatively unaffected (Fig. 3). The chief effect is a selective increase of the unfolding rate constant for the monomeric species. This effect seems fully consistent with the structural location of the mutated side chain (Fig. 2B). By hydrogen bonding to the backbones of Asp⁹² and Gly⁹³, the carboxylate of Asp⁹⁰ supports the tight turn of loop V between β_5 and β_6 . Loop V forms also the closest contact point between the edge strands 5 and 6 at the open side of the SOD1 half-barrel (Fig. 2B), and contains several ALS-provoking mutations, *i.e.* D90A/V and G93A/C/D/R/S/V. The higher destabilization observed for D90V could be that the substitution to Val, in addition to truncate the carboxylate hydrogen bonds, leads to steric conflicts at the protein surface. Another contributing factor could be the “inverse hydrophobic effect” where substitution of a solvent-exposed hydrophobic side chain to a polar ditto increases protein stability by destabilization of the unfolded state (37).

Again, the results provide weak indications that the extent of mutational perturbation scales with the pathogenic capacity. The average survival times after onset are 8 years ($n = 15$) and 2.7 ($n = 3$) for patients heterozygous for the D90A and D90V mutations, respectively. In addition to this, the disease penetrance and progression of D90A is notoriously variable with both rapidly progressing heterozygote patients and slowly progressing homozygote cases depending on the geographical region (14, 31, 38).

D101G and D101N—Interestingly, the structural impact of D101G and D101N is quite different. D101G leads to a pronounced destabilization of both the monomeric and dimeric apoSOD1 species whereas the mutation D101N appears perfectly wild-type like or even somewhat stabilizing. The significance, or cause, of this stabilization that is manifested in slight decrease of k_u for the apoSOD1 monomer is yet not clear. In the folded state, Asp¹⁰¹ (β_6) forms a seemingly critical and strictly conserved salt link with Arg⁷⁹ (end of loop IV, close to β_5). This salt link, which is also relatively embedded by the surrounding structure, could serve two roles. First, it bridges the edge strands β_5 and β_6 at their widest separation at the open end of the SOD1 half barrel (Fig. 2F). Second, it provides some specificity to the conformation of loop IV by a hydrogen-bond network that anchors Arg⁷⁹ with the backbones of Pro⁷⁴, His⁸⁰, Val⁸¹, and Asp⁸³ (Fig. 2F). From a structural perspective, the iso-steric substitution D101N would then be relatively benign by affecting only the electrostatic component of the interaction with the Arg⁷⁹ head group, whereas the cavity-creating mutation D101G would be worse by also truncating several Van der Waals contacts. In the latter case, the perturbation seems further to propagate to the dimer interface. Despite these differences both D101N and D101G yield remarkably rapid disease progressions with survival times after first diagnosis of 2.3 years ($n = 17$) and 1.9 years ($n = 3$), respectively. By comparison with the structurally related mutation D90A, D101N seems thus to be disproportionately aggressive. In common with heterozygous D90A, however, D101N displays incomplete disease penetrance (39).

N139K and N139D—Our final choice of ALS-associated mutations fell on N139K and N139D that target a highly conserved position in the electrostatic loop VII. Of particular interest is that this pair of mutations involves both an increase and decrease of the net repulsive charge. The mutations affect mainly the two hydrogen bonds to Asn¹³¹ that runs along the small E133–T137 helix at the very surface of loop VII (Fig. 2C). Consistent with the surface-exposed position of Asn¹³⁹, the mutations have no detectable effect on either the folding or stability of the apoSOD1 molecule (Fig. 4). Similar to D76Y, however, N139K displays only half of the normal SOD1 activity in lymphoblasts (40). A plausible explanation for this effect is that the mutation reduces the levels of mature SOD1 by introducing an additional protease cleavage site in loop VII, or makes the protein more susceptible to proteolytic cleavage by curbing the hydrogen bonds to Asn¹³¹. The obvious, yet unresolved, question is then if N139K, like D76Y, has a slow disease progression. It will also be interesting to find out if N139D has even lower disease impact (or is even a mistaken polymorphic variant), as this mutation does not introduce a new protease site,

³ P. M. Andersen, unpublished data.

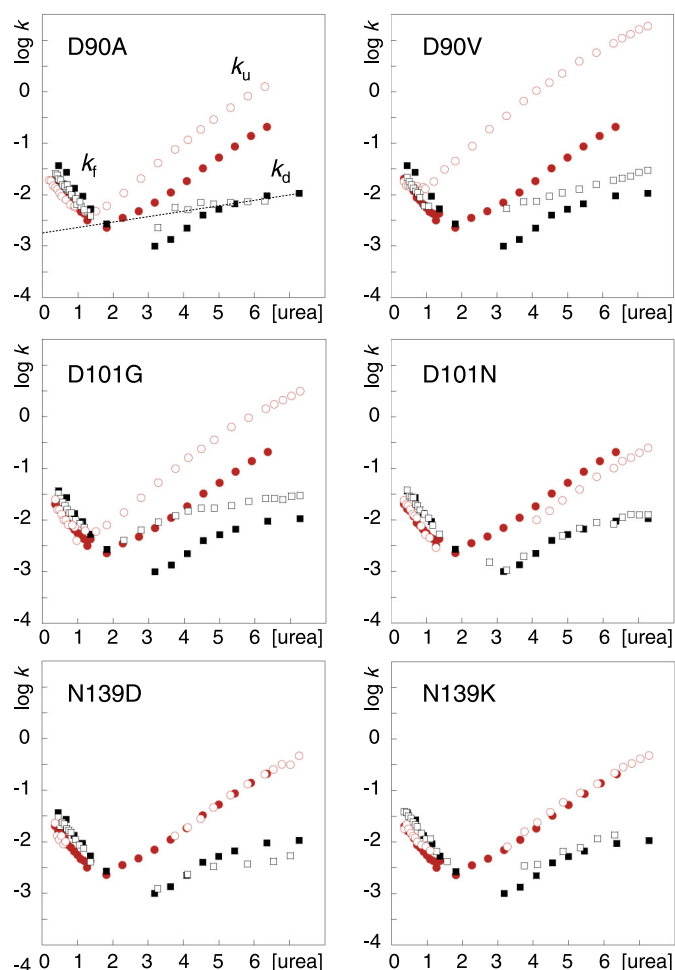


FIGURE 4. Chevron plots of D90V, D90A (17), D101N, D101G, N139D, and N139K. apoSOD^{pWt} (●), apoSOD^{mut} (○), apoSOD^{dimer} (■) and apoSOD^{mut} (□). The units of [urea] and k are M and s^{-1} , respectively.

has smaller effect on the hydrogen bonding and also increases the net repulsive charge. For comparison, the related mutation N139H has a notably aggressive progression with a survival time of 3 years ($n = 6$) after first diagnosis but still incomplete penetrance (41). If the effect on stability of N139H is as modest as observed for N139K/D, it categorizes as a putative outlier together with D101N.

Critical Assessment of the Protein Aggregation Model—As a minimalist test of the role of SOD1 aggregation in the ALS mechanism, several studies have searched for simple relations between the physical-chemical properties of the ALS-provoking mutations and their associated clinical manifestation (12, 13, 17, 42). One of the most controversial result from these examinations is the indication of a correlation between the extent of SOD1 stability loss and patient survival time after first diagnosis (8–10). The key problem is that insufficient clinical statistics and an intrinsic variability of the disease progression for many ALS mutations render the data inconclusive. Nevertheless, the verification or falsification of such a correlation constitute a critical test of the aggregation hypothesis, and of the relation between ALS and other neurodegenerative diseases where the toxic gain-of-function has been linked to the aggregation propensity of the precursor proteins (15, 43). To illustrate the scatter of the data, an extended version of the stability-progression plot is shown in Fig. 5. In addition to the new data presented in Table 1, the plot includes additional new data for the mutations L38V, S105L, G114A, and L144S (supplemental Table S1) and data from Refs. 19–23. To allow comparison of urea- and thermal-denaturation experiments, the measured stabilities have been normalized as relative SOD1 destabilization ($\Delta\Delta G^{\text{norm}}$). Taken together, the data show no compelling correlation between $\Delta\Delta G^{\text{norm}}$ and disease progression. Even so, a basic pattern emerges in the distribution of data

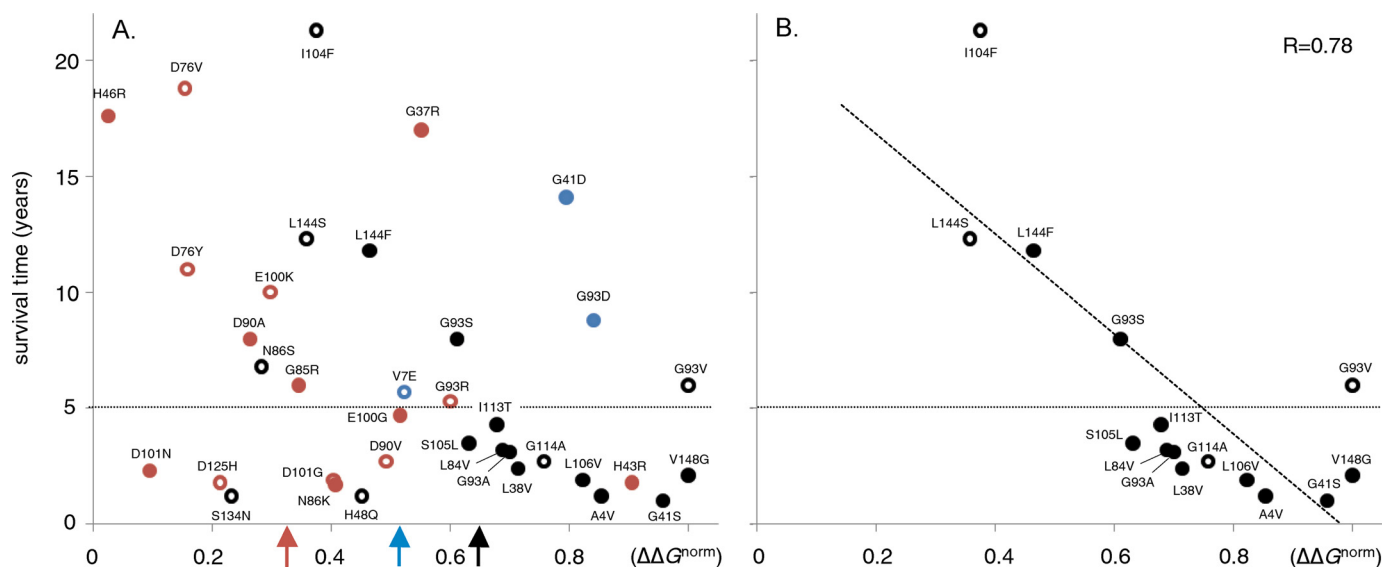


FIGURE 5. Plots of patient survival time after first diagnosis versus the stability loss of the ALS-provoking SOD1 mutation ($\Delta\Delta G^{\text{norm}}$). Data are from supplemental Table S1. The dotted lines are tentatively set at 5 years to denote the division between short and medium/long survival times. A, plot of the combined data set, showing the mutations that do not affect charges (black), the mutations that decrease the net repulsive charge (red), and the mutations that increase the net repulsive charge (blue). Solid circles denotes $n > 5$ and open circles $n < 5$ (supplemental Table S1). The red, blue, and black arrows show the average $\Delta\Delta G^{\text{norm}}$ of each mutant group. B, subset of mutations that do not affect charges, hydrogen bonds or directly affect metal binding, yields an apparent correlation of $r = 0.78$. Although the statistical significance of these patterns are not yet clear, it is interesting to note that they comply with a reductionist disease model based on the propensity of SOD1 aggregation.

ALS Mutations Targeting SOD1 Hydrogen Bonds

points upon separation of the mutations according to their effect on the SOD1 charges. Upon elimination of the few ALS mutations that increase the net repulsive charge (blue) it is apparent that the most destabilizing mutations show invariably very short survival times: the scatter in survival times for severely destabilizing mutations seems due to deviating charge replacements. The pattern complies nicely with assumption that increased repulsive charge is a protective factor because it opposes aggregation. These substitutions are also underrepresented in the list of ALS-associated mutations and have a slower than average disease progression, *i.e.* 7 years (blue only) *versus* 5.6 years (the rest) (Fig. 5). *Vice versa*, it is apparent that the mutations that decrease the net repulsive charge (red) are biased to lower values of $\Delta\Delta G^{\text{norm}}$, and that the mutations at this end of the spectrum have more variable survival times. One interpretation of this $\Delta\Delta G^{\text{norm}}$ bias is that the ALS mutations

that decrease the net repulsive charge require on the whole a lower stability cutback to reach the aggregation/toxic threshold. Another, more reductionist, explanation is that charge substitutions more frequently occur at the SOD1 surface where the stability losses are generally smaller. Neither of these explanations, however, provides a satisfactory answer to the increased scatter in disease progression. From the aggregation model it is expected that diminished repulsive charge would overall accelerate the disease, but clearly there are both short and long survival times at overlapping values of $\Delta\Delta G^{\text{norm}}$. The question then arises, are the low- $\Delta\Delta G^{\text{norm}}$ mutations with rapid disease progression outliers? If so, these mutations would involve specific/local effects that add to the global parameters $\Delta\Delta G^{\text{norm}}$ and net repulsive charge. One such effect could be the rupture of conserved hydrogen bonds, or electrostatic networks, on the SOD1 surface that need not to affect the global protein stability (Figs. 3 and 4 and Table 1). From this perspective it is interesting to see that the most obvious outliers in Fig. 5 are all targeting conserved hydrogen bonds and salt links in loops IV and VII, *i.e.* N86K, D101N, D101G, D125H, and S134N. Stretching the point even further, elimination of all the mutants associated with the rupture of surface hydrogen bonds and charge alterations leaves a seemingly correlated data set with $r = 0.78$. The common feature of these remaining mutations is that they represent pure perturbations of the global SOD1 stability, either induced by steric obstruction or by the truncation of native, hydrophobic contacts. Even so, it is important to realize that, upon selective analysis of complex data sets, correlative patterns will eventually emerge by chance. Therefore we refrain from judging the significance of these data patterns: we just note that they can, tentatively, be rationalized according to a simplistic aggregation model.

Distribution of Charge Alterations among Patients with Rapid and Slow ALS Progression—Finally, as an independent test of how the ALS progression responds to the molecular identity of the SOD1 mutations, we examined the distribution of charge mutations among patient with short (<5 years, $n = 546$) and long (≥ 5 years, $n = 233$) survival time after first diagnosis. The reference set of 75 confirmed ALS-associated SOD1 mutations with progression data were obtained from Refs. 13, 14 and is listed in supplemental Table S2. The result shows that the proportion of charge-altering mutations is substantially higher in the ALS cases with slow disease progression. For survival times below and above 5 years the fractions of charge alterations are 30 and 50%, respectively (Fig. 6A). Looking at the absolute numbers of mutations, however, the charge alterations display no preference for long or short survival times (14 *versus* 14). The bias arises because

TABLE 1

The kinetic parameters for the monomeric and dimeric apoSOD1 species analyzed in this study

$\log k_f^{\text{H}_2\text{O}}$ and $\log k_u^{\text{H}_2\text{O}}$ are obtained from chevron data according to Eq. 2, $\Delta\Delta G_{\text{D-M}}$ is derived from Eq. 3 and $\Delta\Delta G_{\text{D-M}_2}$ is derived from Eq. 4. $\log k_d^{5.8\text{M}}$ shows the dimer-dissociation rate constant at 5.8 M urea.

apoSOD _{mono}	$\log k_f^{\text{H}_2\text{O}}$	$\log k_u^{\text{H}_2\text{O}}$	$\Delta\Delta G_{\text{D-M}}$	apoSOD _{dimer}	$\log k_d^{5.8\text{M}}$	$\Delta\Delta G_{\text{D-M}_2}$
			kcal/mol			
			kcal/mol			
pWT	-1.36	-3.59	-	pWT	-2.13	-
L38V ^a	-1.62 ^a	-2.19	2.25	L38V	-1.4	3.24
D76V	-1.24	-3.32	0.20	D76V	-2.23	0.06
D76Y	-1.48	-3.55	0.21	D76Y	-2.22	0.09
N86D	-1.33	-3.59	-0.05	N86D	-1.4	0.94
N86K	-1.46 ^a	-3.20	0.66	N86K	-1.58	1.41
N86S	-1.25	-3.53	-0.07	N86S	-1.75	0.45
D90V	-1.18 ^a	-2.37	1.41	D90V	-1.8	1.85
D90A ^b	-1.54	-3.28	0.66	D90A	-2.14	0.65
D101G	-1.44 ^a	-3.14	0.72	D101G	-1.64	1.39
D101N	-1.31 ^a	-4.12	-0.80	D101N	-2.09	-0.75
S105L	-1.20 ^a	-2.10	1.81	S105L	-1.55	2.60
G114A	-1.98 ^a	-2.52	2.33	G114A	-1.44	3.27
N139D	-1.43 ^a	-3.64	0.01	N139D	-2.4	-0.36
N139K	-1.30 ^a	-3.48	0.06	N139K	-2.0	0.24
L144S	-1.39	-3.47	0.20	L144S	-1.49	1.07

^a Because of short refolding limbs of the chevron data, m_f was locked to the wild-type value of -0.95 (Eq. 2).

^b Ref. 17.

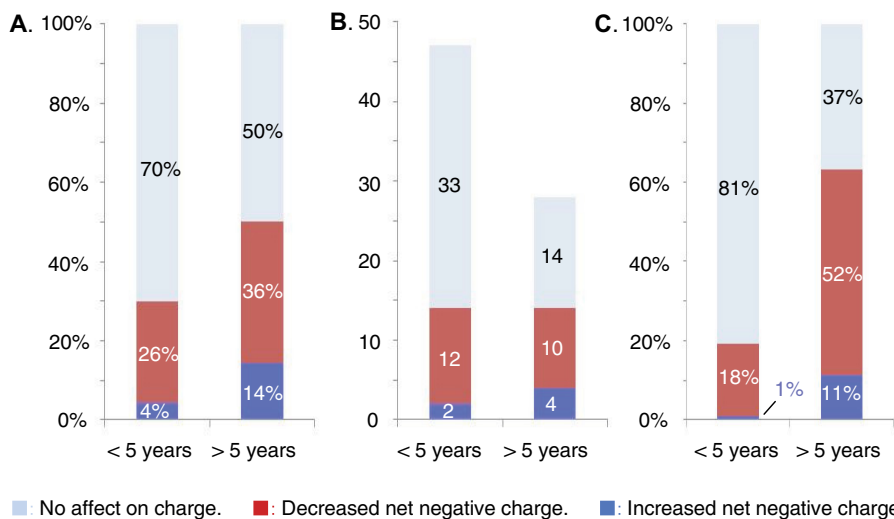


FIGURE 6. The statistical distribution of charge alterations among ALS-associated SOD1 mutations with survival times below and above 5 years after first diagnosis. A, fractions of the total number of different ALS mutations. B, absolute numbers of different ALS mutations. C, fractions of the total number of ALS cases involving SOD1 mutations.

the mutations that not affect charge have predominantly short survival times (33 *versus* 14) (Fig. 6B). Notably, the trend is even more apparent in relation to the total number of ALS cases with SOD1 mutations: charge replacements are found for 63% of the cases with slow disease progression but only for 19% of the cases with rapid progression (Fig. 6C). Rapid ALS progression is thus rarely associated with charge alterations. This observation, however, does not imply that charge alterations are protective as suggested in Ref. 13. Rather, these mutations seem to represent relatively mild perturbations, unless they involve substantial stability losses or the truncation of key hydrogen bonds (see Fig. 5). Loss of a single negative charge appears sufficient for ALS penetrance (5). The only molecular factor that is yet indicated to oppose ALS is the addition of extra negative charges to the SOD1 molecule (or loss of positive dittos), because these are under-represented among the disease-provoking mutations (5). Such a suppressing role of augmented repulsive charge is also in line the overall slow disease progression for the severely destabilizing mutations G41D ($n = 15$) and G93D ($n = 7$), where the similarly destabilized, but charge neutral, mutations G41S ($n = 16$) and G93A ($n = 16$) have clearly shorter survival times (Fig. 5). On this basis we conclude that, even if the relation between SOD1 stability and ALS progression is yet uncertain, there is a statistically significant overrepresentation of charge alterations among the ALS-associated mutations with survival times >5 years after first diagnosis.

Concluding Remarks—Naturally evolved proteins seem to employ two basic strategies to remain soluble in the crowded interior of the cell: (i) they protect potentially sticky epitopes in their folded interior (44), and (ii) they use charges to obstruct aggregation by negative design (1–3, 45, 46). If ALS is indeed linked to the precursor protein's intrinsic propensity to aggregate (47, 48), as is implicated for other protein-deposition diseases (44, 49–51), we would also expect a coupling between the outbreak/progression of the disease and the physical-chemical properties of the ALS-provoking SOD1 mutations. In the simplest case, with decreased protein stability because this increases the amount of unfolded and partly unfolded SOD1 material. The most apparent outliers in such a relation, if it exists, are the severely destabilized mutations with slow ALS progression and, *vice versa*, the marginally destabilized mutations with fast progression (Fig. 5). In the first set we find G41D and G93D that both increase the net repulsive charge. According to elementary electrostatic considerations these mutations are expected to suppress the aggregation propensity and, consistently, G41D is observed to aggregate less than the charge neutral but equally destabilizing mutation G41S (17). In the second set, intriguingly, we find mutations that target conserved hydrogen bond networks or electrostatic contacts at the surface of loops IV and VII. The mechanistic role of these mutations seems to extend beyond stability and charge of the apo-protein. One possibility is that they make the loop geometry more susceptible to distortions and breathing motions (35). This is perfectly consistent with the computational observation that altered loop dynamics modulates the aggregation of ALS-associated SOD1 mutants because of long-range communication within the SOD1 structure (52) (35). Such increased breathing of the loops could influence how SOD1 holds onto its

metal cofactors and/or render the structure more sensitive to degradation by proteases. Some of these loop-targeting SOD1 mutations display also decreased protein levels/enzymatic turnover in human erythrocytes (31, 32, 42). Taken together, the results from this study point thus at the possibility that mutational perturbation of hydrogen bonds and electrostatic interactions at the SOD1 surface constitute an independent ALS factor that needs not to be accompanied by alterations of apoSOD1 stability or charge to trigger disease.

REFERENCES

- Otzen, D. E., Kristensen, O., and Oliveberg, M. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9907–9912
- Richardson, J. S., and Richardson, D. C. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2754–2759
- Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C. M. (2003) *Nature* **424**, 805–808
- Alsod (2009) *The ALS Online Database*, <http://alsod.iop.kcl.ac.uk/Als/index.aspx>
- Sandelin, E., Nordlund, A., Andersen, P. M., Marklund, S. S., and Oliveberg, M. (2007) *J. Biol. Chem.* **282**, 21230–21236
- Lindberg, M. J., Normark, J., Holmgren, A., and Oliveberg, M. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15893–15898
- Shaw, B. F., and Valentine, J. S. (2007) *Trends Biochem. Sci.* **32**, 78–85
- Shibata, N., Hirano, A., Kobayashi, M., Sasaki, S., Kato, T., Matsumoto, S., Shiozawa, Z., Komori, T., Ikemoto, A., and Umahara, T. (1994) *Neurosci. Lett* **179**, 149–152
- Shibata, N., Hirano, A., Kobayashi, M., Asayama, K., Umahara, T., Komori, T., and Ikemoto, A. (1993) *Acta Histochem. Cytochem.* **26**, 619–622
- Watanabe, M., Dykes-Hoberg, M., Culotta, V. C., Price, D. L., Wong, P. C., and Rothstein, J. D. (2001) *Neurobiol. Dis.* **8**, 933–941
- Jonsson, P. A., Ernhill, K., Andersen, P. M., Bergemalm, D., Brännström, T., Gredal, O., Nilsson, P., and Marklund, S. L. (2004) *Brain* **127**, 73–88
- Lindberg, M. J., Tibell, L., and Oliveberg, M. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16607–16612
- Wang, Q., Johnson, J. L., Agar, N. Y., and Agar, J. N. (2008) *Plos Biology* **6**, 1508–1526
- Andersen, P. M. (2006) *Curr. Neurol. Neurosci. Rep.* **6**, 37–46
- Chen, S. M., Ferrone, F. A., and Wetzel, R. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11884–11889
- Landles, C., and Bates, G. P. (2004) *EMBO Rep.* **5**, 958–963
- Lindberg, M. J., Bystrom, R., Boknas, N., Andersen, P. M., and Oliveberg, M. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9754–9759
- Fersht, A. (1999) *Structure and Mechanism in Protein Science: a Guide to Enzyme Catalysis and Protein Folding*, W.H. Freeman and Company, New York
- Stathopoulos, P. B., Rumpfledt, J. A., Scholz, G. A., Irani, R. A., Frey, H. E., Hallewell, R. A., Lepock, J. R., and Meiering, E. M. (2003) *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7021–7026
- Stathopoulos, P. B., R. J., Karbassi, F., Siddall, C. A., Lepock, J. R., and Meiering, E. M. (2006) *J. Biol. Chem.* **281**, 6184–6193
- Vassall, K. A., Stathopoulos, P. B., Rumpfledt, J. A. O., Lepock, J. R., and Meiering, E. M. (2006) *Biochemistry* **45**, 7366–7379
- Rodriguez, J. A., Shaw, B. F., Durazo, A., Sohn, S. H., Doucette, P. A., Nersissian, A. M., Faull, K. F., Eggers, D. K., Tiwari, A., Hayward, L. J., et al. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10516–10521
- Furukawa, Y., and O'Halloran, T. V. (2005) *J. Biol. Chem.* **280**, 17266–17274
- Ferraroni, M., Rypniewski, W., Wilson, K. S., Viezzoli, M. S., Banci, L., Bertini, I., and Mangani, S. (1999) *J. Mol. Biol.* **288**, 413–426
- Levy, Y., Wolynes, P. G., and Onuchic, J. N. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101**, 511–516
- Svensson, A. K. E., Bilsel, O., Kondrashkina, E., Zitzewitz, J. A., and Matthews, C. R. (2006) *J. Mol. Biol.* **364**, 1084–1102
- Tiwari, A., and Hayward, L. J. (2005) *Neurodegenerative Dis.* **2**, 115–127
- Strange, R. W., Antonyuk, S., Hough, M. A., Doucette, P. A., Rodriguez,

ALS Mutations Targeting SOD1 Hydrogen Bonds

- J. A., Hart, P. J., Hayward, L. J., Valentine, J. S., and Hasnain, S. S. (2003) *J. Mol. Biol.* **328**, 877–891
29. DiDonato, M., Craig, L., Huff, M. E., Thayer, M. M., Cardoso, R. M., Kassmann, C. J., Lo, T. P., Bruns, C. K., Powers, E. T., Kelly, J. W., *et al.* (2003) *J. Mol. Biol.* **332**, 601–615
30. Hubbard, S. J., Beynon, R. J., and Thornton, J. M. (1998) *Protein Eng.* **11**, 349–359
31. Andersen, P. M., Nilsson, P., Keranen, M. L., Forsgren, L., Hagglund, J., Karlsborg, M., Ronnevi, L. O., Gredal, O., and Marklund, S. L. (1997) *Brain* **120**, 1723–1737
32. Orrell, R. W., Habgood, J. J., Gardiner, I., King, A. W., Bowe, F. A., Hallewell, R. A., Marklund, S. L., Greenwood, J., Lane, R. J., and deBeleroche, J. (1997) *Neurology* **48**, 746–751
33. Tiwari, A., and Hayward, L. J. (2003) *J. Biol. Chem.* **278**, 5984–5992
34. Ding, F., and Dokholyan, N. V. (2008) *Proc. Natl. Acad. Sci. U. S. A.* **105**, 19696–19701
35. Khare, S. D., and Dokholyan, N. V. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3147–3152
36. Ferguson, N., Pires, J. R., Toepert, F., Johnson, C. M., Pan, Y. P., Volkmer-Engert, R., Schneider-Mergener, J., Daggett, V., Oschkinat, H., and Fersht, A. (2001) *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13008–13013
37. Munoz, V., Lopez, E. M., Jager, M., and Serrano, L. (1994) *Biochemistry* **33**, 5858–5866
38. Andersen, P. M., Sims, K. B., Xin, W. W., Kiely, R., O'Neill, G., Ravits, J., Pioro, E., Harati, Y., Brower, R. D., Levine, J. S., *et al.* (2003) *Amyotroph Lateral Scler Other Motor Neuron Disord* **4**, 62–73
39. Jones, C. T., Shaw, P. J., Chari, G., and Brock, D. J. H. (1994) *Mol. Cell Probes* **8**, 329–330
40. Pramatarova, A., Figlewicz, D. A., Krizus, A., Han, F. Y., Ceballos-Picot, I., Nicole, A., Dib, M., Meiningner, V., Brown, R. H., and Rouleau, G. A. (1995) *Am. J. Hum. Genet.* **56**, 592–596
41. Nogales-Gadea, G., Garcia-Arumi, E., Andreu, A. L., Cervera, C., and Gamez, J. (2004) *J. Neurol. Sci.* **219**, 1–6
42. Sato, T., Nakanishi, T., Yamamoto, Y., Andersen, P. M., Ogawa, Y., Fukada, K., Zhou, Z., Aoike, F., Sugai, F., Nagano, S., *et al.* (2005) *Neurology* **65**, 1954–1957
43. Stefani, M., and Dobson, C. M. (2003) *J. Mol. Med.* **81**, 678–699
44. Chiti, F., and Dobson, C. M. (2006) *Annu. Rev. Biochem.* **75**, 333–366
45. Thirumalai, D., Klimov, D. K., and Dima, R. I. (2003) *Curr. Opin. Struct. Biol.* **13**, 146–159
46. Rousseau, F., Schymkowitz, J., and Serrano, L. (2006) *Curr. Opin. Struct. Biol.* **16**, 118–126
47. Furukawa, Y., Kaneko, K., Yamanaka, K., O'Halloran, T. V., and Nukina, N. (2008) *J. Biol. Chem.* **283**, 24167–24176
48. Shaw, B. F., Lelie, H. L., Durazo, A., Nersissian, A. M., Xu, G., Chan, P. K., Gralla, E. B., Tiwari, A., Hayward, L. J., Borchelt, D. R., Valentine, J. S., and Whitelegge, J. P. (2008) *J. Biol. Chem.* **283**, 8340–8350
49. Balch, W. E., Morimoto, R. I., Dillin, A., and Kelly, J. W. (2008) *Science* **319**, 916–919
50. Luheshi, L. M., Crowther, D. C., and Dobson, C. M. (2008) *Curr. Opin. Chem. Biol.* **12**, 25–31
51. Luheshi, L. M., Tartaglia, G. G., Brorsson, A. C., Pawar, A. P., Watson, I. E., Chiti, F., Vendruscolo, M., Lomas, D. A., Dobson, C. M., and Crowther, D. C. (2007) *Plos Biol.* **5**, 2493–2500
52. Banci, L., Bertini, I., Boca, M., Calderone, V., Cantini, F., Girotto, S., and Vieru, M. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6980–6985
53. Homologene (2009) in <http://www.ncbi.nlm.nih.gov/homologene>
54. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) *Genome Res.* **14**, 1188–1190
55. Notredame, C., Higgins, D. G., and Heringa, J. (2000) *J. Mol. Biol.* **302**, 205–217