Surface charges of proteins have in several cases been found to function as “structural gatekeepers,” which avoid unwanted interactions by negative design, for example, in the control of protein aggregation and binding. The question is then if side-chain charges, due to their desolvation penalties, play a corresponding role in protein folding by avoiding competing, misfolded traps? To find out, we removed all 32 side-chain charges from the 101-residue protein S6 from *Thermus thermophilus*. The results show that the charge-depleted S6 variant not only retains its native structure and cooperative folding transition, but folds also faster than the wild-type protein. In addition, charge removal unleashes pronounced aggregation on longer timescales. S6 provides thus an example where the bias toward native contacts of a naturally evolved protein sequence is independent of charges, and point at a fundamental difference in the codes for folding and intermolecular interaction: specificity in folding is governed primarily by hydrophobic packing and hydrogen bonding, whereas solubility and binding relies critically on the interplay of side-chain charges.

**Results**

**Design of a Charge-free Protein.** Charge removal was done in two steps (Fig. 1). First, all K and R side chains in wild-type S6 were mutated to S. The resulting supercharged variant S6 expressed in soluble form with good yields in crystalline form. Our conjecture from these data is that the profusion of charges scattered in the sequences of natural proteins are not required for folding per se, but play their major role in solubility, recognition, and biological function.

**Charge-depleted S6 Maintains Native-like Solution Structure.** Wild-type S6 has a fixed, tertiary-ordered structure in solution, which is indistinguishable from that in crystals (23). In this study, we examine at more general level the role of side-chain charges in protein folding and aggregation by removing them completely. S6 is normally rich in charges and carries 16 negatively and 16 positively charged side chains, comprising 32% of its sequence content. By a combination of protein engineering and lowered pH we produced a protein that is altogether non-charged, save the positive N-terminal (S6). The results show that complete charge removal, if anything, favors the folding process: S6 not only maintains a classically v-shaped chevron plot, but also folds faster than the wild-type protein. On longer timescales, however, the protein starts to aggregate, both with native and denatured S6 as starting material. Our conjecture from these data is that the profusion of charges scattered in the sequences of natural proteins are not required for folding per se, but play their major role in solubility, recognition, and biological function.
examine the conservation of secondary structure we plotted the
(Fig. S3). Wild-type S6 is a two-state process with two competing
pathways, the bias barrier (28). The intersect of log the kinetic prefactor; i.e., the rate constant for jumping down the
way (26), manifested in a v-shaped chevron plot characteristic of
17 negative charges at low pH (Table 1) is then unlikely to have
neutrality, and the structurally ordered permutant S6 is mostly
assigned was obtained unambiguously for 84 of the 101 resi-
due to electrostatic repulsion within the S6 structure (24). Even so, C1–C1
secondary-chemical-shift analysis indicate that the secondary-structure
elements remain wild-type like at all positions where the assignments of S6, S6, and S6 overlap (Fig. S2). To further examine the conservation of secondary structure we plotted the
secondary chemical shifts of wild-type S6 versus those of
S6 (Fig. S2). The results yield a linear correlation of r = 0.95, which closely resembles that between wild-type S6 and the structurally ordered permutant S6 (23, 24) (Fig. S2), and supports the conclusion that K and R removal does not significantly alter the S6 structure. Neutralization of the remaining 17 negative charges at low pH (Table 1) is then unlikely to have
any further structural effects: if anything, the complete alleviation of electrostatic repulsion within the S6 molecule is expected to
be beneficial. Consistently, the charge-depleted S6 structure
yields highly dispersed HSQC spectra under carefully tuned condi-
tions at pH 1 (Fig. 2), indicating a homogeneous, tertiary-
ordered population. However, the disappearance of signal due to
aggregation on longer timescale has so far precluded detailed structural assignment of the charge-depleted protein.

Charges are not Required for Cooperative Folding of S6. Folding of
S6 is a two-state process with two competing pathways, the bias
between which can be altered by circular permutation (24, 25)
(Fig. S3). Wild-type S6 unfolds and refolds through at least one of these pathways (26), manifested in a v-shaped chevron plot characteristic of a cooperative transition between the denatured (D) and native (N) states over a single transition-state (‡) (25–27) where kD and kU are the refolding- and unfolding-rate constants, and kDref is the kinetic prefactor; i.e., the rate constant for jumping down the barrier (28). The intersect of log kD and log kU in the S6 chevron plot gives a transition midpoint (MP) at 3.4 M GdmCl
and an extrapolated stability of ΔG29° C = 8.2 kcal/mol (Fig. 3, Eq. 4, Table 1). Upon mutual removal of all positively charged side chains, the transition midpoint decreases to 1.7 M GdmCl (Fig. 3, Table 1) and the chevron changes yield a global φ-value of 0.20; i.e., the mutation to S6 has the largest impact on log k0 [Eq. 5]. This means that global removal of alkaline residues destabilizes N more than D. Considering the high ionic strength under the experimental conditions; i.e., [GdmCl] > 0.4 M, this destabilization is not dominated by long-range electrostatic repulsion but seems rather to arise from mutant-induced contact losses and/or changes in solvent interactions. At low ionic strength with added charge-charge repulsion the destabilization of S6 becomes even larger, leading to mixed populations of D and N (Fig. S4). Finally, to obtain the folding kinetics of the charge-depleted species S6, we adjusted the refolding- and unfolding buffers to pH 2.3, to assure full protonation of the acidic side chains in the reaction chamber of the stopped-flow appara-
tus. Two notable features are revealed. First, the charge-depleted protein maintains a v-shaped chevron plot and apparent two-state behavior in the absence of side-chain charges. Second, charge de-
pletion increases both kD and kU; at the transition midpoint S6 folds about 300 times faster than wild-type S6 (Fig. 3, Table 1). The corresponding data for wild-type S6 at pH 2.3 are shown in Fig. S5. As the acceleration upon charge depletion is not accompanied by a correspondingly large change in protein stability (Fig. 4, Table 1), it seems to arise from either a selective stabilisation of the transition-state ensemble (‡) or an increase of the folding prefactor; i.e., kDref in Scheme 1 (10, 28). The chevron plots of S6 (Figs. 3 and 4) reveal also slight reductions of mK and mU [Eqs. 2–3] and S6 displays a small downward curvature at high [GdmCl] (Fig. 3). Similar m-value changes and curvatures are seen upon point mutation of the wild-type and circularly permuted S6 (29), and are generally attributed to transition-state shifts (30) (Fig. 4, Fig. S3) or ground-state fraying (27). In the present study, however, it is also possible that the m-value changes stem from alterations of the protein’s interaction with the solvent/denaturant molecules or changes of the compact-
ness of the denatured ensemble as described below (Fig. 4).

Fig. 1. Charge removal of S6. A. Wild-type S6 comprises 16 negative (red) and 16 positive (blue) charges on the protein surface. B. Supercharged S6 was produced by mutation of all K and R to S. C. Charge-depleted S6 was finally obtained by transferring S6 to pH 2.3 where all negative side-chain charges as well as the C-terminal becomes neutralized by protonation. D. Positions of positively- (blue) and negatively (red) charged side chains in the S6 sequence.

Fig. 2. NMR HSQC spectra of the different charge variants of S6. All spectra show wild-type like dispersion, suggesting that the supercharged S6 and the charge-depleted S6 maintain fixed, three-dimensional structures. Additional NMR evidence for conserved structure of S6 is presented in Fig. S2, whereas detailed structural analysis of S6 has so far been precluded by aggregation on longer timescales (Fig. 5).
the GdmCl concentration at a background of 1.0 M Na$_2$SO$_4$ (Fig. S6). For simplicity, we denote this new phase $k^c$ (Scheme 2). The phenomenon has previously been assigned to a change of rate-limiting step: when the free-energy difference between the collapsed state and the normal $\xi$ reaches a critical value, folding switches to a parallel pathway (5). Since the denaturant and Na$_2$SO$_4$ dependencies of $\log k^c$ are very much weaker than for $\log k_f$, this parallel pathway was suggested to show reconfigurations between collapsed species of similar solvent accessible surface area—a SO$_2^-$-induced detour through compact regions of the conformational space (5). In this study, we observe further that the change of rate-limiting step is coupled to the emergence of several slower refolding phases (Fig. S7). One possibility is that these phases describe the onset of transient aggregation occurring in parallel with the coil collapse, as observed upon removal of the charged aggregation gatekeepers in j2 (20), or stem from a more complex partitioning of trapped species under high misfolding pressure. In keeping with the recent discussion of downhill folding (33, 34), the complex time course of $k_f$ could also reflect a change to stretched-exponential behavior arising from barrierless reconfigurations in an increasingly rugged landscape. The collapse transition would then break down the delicate imbalance between entropy and contact free energy that shape the barrier, promoting noncooperative downhill folding.

Table 1. Kinetic parameters and protein stabilities

<table>
<thead>
<tr>
<th>Parameter</th>
<th>S6$^+$</th>
<th>S6$^{−1}$</th>
<th>S6$^−$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{f}^{M_n}$ $\pm$</td>
<td>2.53 ± 0.03</td>
<td>1.48 ± 0.03</td>
<td>3.28 ± 0.06</td>
</tr>
<tr>
<td>$m^{*}$ (M$^{−1}$)</td>
<td>−1.22 ± 0.01</td>
<td>−0.82 ± 0.04</td>
<td>−1.10 ± 0.04</td>
</tr>
<tr>
<td>$k_{f}^{M_n}$ $\pm$</td>
<td>−3.51 ± 0.08</td>
<td>−0.64 ± 0.05</td>
<td>0.35 ± 0.04</td>
</tr>
<tr>
<td>$m^{*}$ (M$^{−1}$)</td>
<td>0.54 ± 0.02</td>
<td>0.43 ± 0.01</td>
<td>0.27 ± 0.01</td>
</tr>
<tr>
<td>MP$^+$ (M)</td>
<td>3.44 ± 0.03</td>
<td>1.71 ± 0.05</td>
<td>2.15 ± 0.04</td>
</tr>
<tr>
<td>MP$^2$ ( M$^{−1}$)</td>
<td>1.76 ± 0.02</td>
<td>1.24 ± 0.04</td>
<td>1.36 ± 0.04</td>
</tr>
<tr>
<td>$k_{f}^{M_n}$ $\pm$</td>
<td>6.04 ± 0.09</td>
<td>2.12 ± 0.06</td>
<td>2.93 ± 0.07</td>
</tr>
<tr>
<td>$\Delta G^{M_n}_{f}^{\circ}$ (kcal/mol)</td>
<td>8.24 ± 0.12</td>
<td>2.89 ± 0.08</td>
<td>3.99 ± 0.09</td>
</tr>
</tbody>
</table>

*Derived from chevron data according to Eq. 3. Data for S6$^{−1}$ is derived from the v-shaped regime below [GdmCl] = 5 M.

1Transition midpoint derived from the intersect between log $k_f$ and log $k_{f}^{M_n}$ according to Eq. 2.

2Calculated from Eq. 2.

3Calculated from $\Delta G^{M_n}_{f}^{\circ} = −2.3RT(\log k_{f}^{M_n} − \log k_{f}^{M_n})$ [Eq. 4].

Elimination of 33 charges represents, after all, a considerable change in the chemical properties of the polypeptide chain. Even so, it is clear that the charge depletion of S6 has little effect on the principal features of the folding process, if anything the protein seems to fold faster without charges.

Charge-Depletion Affects Collapse Propensity. To determine how charge removal affects the misfolding propensity, we measured and compared the refolding kinetics of S6$^{+}$, S6$^{−1}$, and S6$^{−}$ at increasing concentrations of Na$_2$SO$_4$. Titration with stabilizing SO$_2^−$ ions has previously been found to induce misfolding of S6, accompanied by a characteristic retardation of $k_f$ (5). The rationale behind this experiment is that titration with cosmotropic SO$_2^−$ ions gradually increases the contact free energies to a point where misfolding start to retard $k_f$ (5); i.e., the frustration in the folding-energy landscape is increased to promote nonnative collapse (12, 31). A minimalistic model for such a collapse is shown in Scheme 2, where $\xi$ is a competing, misfolded state and $k_f^c$ denotes tentatively an alternative folding route to the native state (5). To assure suitable windows for the refolding kinetics the experiments were performed at a background of 0.4 and 1.6 M GdmCl (Fig. 3). At 0.4 M GdmCl, $k_f$ of S6$^{+}$ first increases as $\xi$ is stabilized relative to D by a “reversed” denaturant effect (Fig. 3); i.e., the SO$_2^−$ ions favor compact states by being preferentially excluded from the protein’s hydration shell. Then, around 0.2 M Na$_2$SO$_4$, $k_f$ starts to decrease as misfolding commence (Fig. 3). In an earlier study, we ascribed this misfolding to premature collapse of the coil in the mixing dead time (5), which can slow down folding by either ground-state stabilization or retardation of the diffusive motions. Consistently, the maximum of $k_f$ shifts to higher [Na$_2$SO$_4$] under better solvent conditions at 1.6 M GdmCl (Fig. 3). A similar shift of the $k_f$ maximum is observed for the supercharged S6$^{+}$ (Fig. 3), indicating that the increased negative charge suppresses coil collapse, cf. (32). The very opposite effect is observed for S6$^{−}$: complete charge removal increases slightly the misfolding propensity (Fig. 3). Despite this tendency, the role of charges in smoothening the folding funnel seems overall marginal as this increased collapse propensity does not compromise folding in the absence of Na$_2$SO$_4$ (12).

The Compact Detour. From the S6$^{−1}$ data in Fig. 3, it is evident that log $k_f$ does not continue to decrease with increasing [Na$_2$SO$_4$] but levels off and describes a slight positive slope above 0.8 M Na$_2$SO$_4$. A corresponding change of log $k_f$ is seen upon lowering D. This suggests that a 33 charge reduction of S6 has little effect on the native structure of the protein (33, 34), the complex time course of $k_f$ could also reflect a change to stretched-exponential behavior arising from barrierless reconfigurations in an increasingly rugged landscape.

Discussion

The data in this study demonstrate that the native structure and folding behavior of S6 does not rely on the presence of side-chain charges: the protein displays a swift and cooperative folding transition both with and without side-chain charges (Figs. 2–3). The result concurs with the earlier conclusion by Loladze and Mukhtadze that surface charge–charge interactions are not essential for protein folding, based on thermodynamic analysis of chemically charge-depleted ubiquitin (35). Judging by the accelerated folding kinetics of S6$^{−1}$ (Fig. 3), it can even be said that charges are a burden to protein folding. The origin of this acceleration, however, is not yet clear. One possibility is that side-chain charges restrict the protein’s reconfigurations or ability to collapse (36) by the way they interact with the solvent. Along this line, elimination of charges could speed up folding by increasing the degree of unspecific hydrophobic contacts in the transition-state ensemble, as observed for the $\alpha$-spectrin SH3 domain upon Tyr-Phe
Scheme 2.

Fig. 3. The folding kinetics of the different charge variants of S6 analyzed by stopped-flow mixing. A. The chevron plots of wild-type S6$^{17-17}$, supercharged S6$^{1+11}$ and charge-depleted S6$^{-1}$ are all classically v-shaped, showing that the folding transition remains cooperative and does not rely on side-chain charges. Of particular interest is that complete charge removal even speeds-up folding: at the transition midpoint S6$^{1-}$ folds $>300$ times faster than wild-type S6$^{17-17}$.

B. Na$_2$SO$_4$ titration of the refolding reaction at 0.4 M GdmCl shows that the propensity of the coil to undergo premature collapse in the mixing dead-time slightly increases upon removal of all positively charged side chains. C. At 1.6 M GdmCl, it is seen that charge-depleted S6$^{-1}$ has the highest collapse propensity of the three proteins and also displays a change of rate-limiting step at high [Na$_2$SO$_4$] (log $k_r$). The origin of this change could be the population of an alternative, parallel, folding route to the native state according to Scheme 2.

Fig. 4. Folding free-energy profiles of wild-type S6$^{17-17}$ (blue), supercharged S6$^{1+11}$ (red), and charge-depleted S6$^{-1}$ (black). Barrier heights were calculated from $k_r$ and a prefactor of $10^{4}$ s$^{-1}$ (28). Charge removal leads to faster folding kinetics and an apparent stabilization of the transition-state ensemble ($\dagger$), whereas the native state is destabilized. The positions of D, $\dagger$ and N along the progress coordinate have been scaled according to the m-values in Table 1 and normalized to N.

distinct dependence on charge content (22), and reduction of a protein’s net negative charge by merely one unit is sufficient for triggering fatal neurodegenerative disease (43). These observations go hand-in-hand with the general idea that charges enhance interaction specificity, not primarily by attraction, but because of large penalties for unmatched burial (17, 44, 45). In this context it is interesting to note that the S6 charges do a better job to prevent aggregation than high concentrations of denaturant (Fig. 5). Surveys of structural data banks shows consistently that lysine, which is the most abundant residue on protein surfaces (46), is the most underrepresented amino acid at interfaces between proteins in functional complexes (47, 48). To this end there are also numerous examples where surface charges control protein interactions in a more attractive way; e.g., in protein-protein recognition (49, 50), in binding to membranes (51) and binding of metal ions (52). As a first approximation, we distinguish these charge-controlled interactions, including solubility and spatial organization, as “functional,” since they orchestrate the biological function of the elementary folding units. With this distinction, the ability of S6 to fold without charges raises the question if there is a principal division of amino-acid use in the self-organization of proteins: folding can evidently be driven by hydrophobicity and hydrogen bonding alone, whereas function and intermolecular organization tend to rely critically on the interplay of charges. An advantage of such a separation of “driving forces” could be that, in its pure form, it biases folding and function from being intertwined; i.e., folding and function have chemical space to evolve independently. It is nevertheless inevitable that charges will still have a pronounced influence on their protein scaffolds, be it through favorable ion pairing or as a side effect of conflicting, functional optimization (49). In some cases, charges control even the structural order of entire protein domains. An intriguing example is the “dome-like” active-site envelope of Cu/Zn superoxide dismutase, which is built almost exclusively by polar and charged side chains, pulled together at its center by a single Zn$^{2+}$ ion (53, 54). The metal ion seems here to substitute for a local hydrophobic core and creates effectively a functional subdomain that structures separately and does not interfere with the folding of the...
main hydrophobic core (55). This split architecture and folding behavior of superoxide dismutase lends further support to the conjecture of an underlying, chemical bias in codes for folding and function of proteins. A clue to the question “why are proteins charged” (16) could then be: not for folding of the basic structural domains.

**Materials**

**Mutagenesis, Expression, and Purification.** S6<sup>−17</sup> gene synthesis, codon optimization for overexpression in *E. coli*, subcloning into a pET-3a vector using S’ Nde1 and 3’ BamHI restriction sites, and construct sequencing were performed by GenScript. Transformation into *E. coli* BL21 (DE3) cells was by standard heat-shock procedures. Expression and purification were as previously described for wild-type S6<sup>−17</sup> (27), whereas the supercharged S6<sup>−11</sup>−17 required a modified purification protocol (Supporting Information). Mutagenesis, expression, and purification of S6<sup>−11</sup> were performed by the Protein Analysis Center (Karolinska Institute). Edman degradation showed that S6<sup>−11</sup>−17 lacks the N-terminal methionine present in wild-type S6<sup>−17</sup>.

**NMR Spectroscopy.** HSQC NMR data were obtained at 25°C with protein concentrations ranging from 30 (pH 1) to 500 μM (pH 6.3, 100 mM NaCl), on a Bruker 700 MHz spectrometer (Bruker Avance) equipped with a cryogenically cooled triple resonance probe. Backbone assignment was obtained from a set of standard Bruker 700 MHz spectrometer (Bruker Avance) equipped with a cryogenically cooled triple resonance probe. Backbone assignment was obtained from a set of standard HSQC, HNCA, HN(CO)CA, HNCO, experiments on a 800 MHz Varian (Varian) Spectra were transformed using NMRPipe and analyzed with the program Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, UCSF).

**Kinetic Measurements.** Refolding and unfolding kinetics were monitored at 25°C with PIsStar-180 and SX-18-MV stopped-flow fluorimeters (Applied Photophysics). Excitation was at 280 nm and emission was collected with a 320 nm long-pass filter. Protein concentration after mixing was 1 μM. Buffers were: 50 mM Mes (Sigma-Aldrich) at pH 6.3, 50 mM formate at pH 3.5–4.5 (Scharlau), and at pH ≤ 3.0, the concentration of HCl corresponding to the pH. Between pH 1.3 and pH 3.0, NaCl (VWR) was added to achieve a final ionic strength of 50 mM unless otherwise stated. Ultrapure guanidinium hydrochloride (AppliChem) and proanalysis Na<sub>2</sub>SO<sub>4</sub> (Merck) were used in the denaturation experiments.