ELSEVIER

Biophysical Chemistry 110 (2004) 239–248

Biophysical Chemistry

www.elsevier.com/locate/bpc

Proteins with simplified hydrophobic cores compared to other packing mutants

Junmei Chen[1], Zhiqiang Lu[2], Joshua Sakon, Wesley E. Stites*

Department of Chemistry and Biochemistry, University of Arkansas, CHEM 101, Fayetteville, AR 72701-1201, USA

Received in revised form 6 February 2004; accepted 27 February 2004
Available online 27 April 2004

Abstract

Efforts to design proteins with greatly reduced sequence diversity have often resulted in proteins with so-called molten globule properties. Substitutions were made at six neighboring sites in the major hydrophobic core of staphylococcal nuclease to create variants with all leucine, all isoleucine or all valine at these sites. The mutant proteins with simplified cores constructed here are quite unstable and have poorly packed cores, attested to by interaction energies. Eight related mutants with greater sequence diversity were also constructed. Comparison to these mutants and 159 other permutations of these 3 aliphatic side chains at these same 6 sites previously constructed shows that the simplified cores are not unusual in their stabilities or interaction energies. Further, crystal structures of the two mutants with the worst packing, as measured by interaction energies, showed no unusual disorder in the core. Therefore, reduction of sequence diversity is not necessarily incompatible with a single stable native structure. Other factors must also contribute to previous protein design failures.

Keywords: Hydrophobic packing; Protein structure; Side chain conformation

A tightly packed hydrophobic core, minimizing cavities and maximizing favorable van der Waals contacts, is very important for protein structure and stability [1–6]. Richards [7] pointed out that protein cores are about as tightly packed as an organic crystal. This is rather remarkable given the complexity and lack of symmetry in a protein interior compared to a crystal of a single type of small molecule. Naively, one might expect that a simple core with a limited number of side chain types, perhaps even just a single type of side chain, might more readily find a good packing solution than a diverse mixture of many different hydrophobic side chains.

Indeed, many initial attempts to design proteins used fairly simple hydrophobic cores utilizing a small subset of amino acids. Such designs often formed a compact structure similar to so-called molten globule intermediates found upon unfolding in some naturally occurring proteins [8–11]. Molten globules have the characteristics of both folded and unfolded proteins. They seem to have α-helix and β-sheet secondary structures like native proteins, but they have no well-

* Corresponding author. Tel.: +1-479-575-7478; fax: +1-479-575-4049.
E-mail address: wstites@uark.edu (W.E. Stites).
[1] Current address: N1302 Alkek Research Building, One Baylor Plaza, Baylor College of Medicine, Houston, TX 77030.
[2] Current address: Johns Hopkins School of Public Health, MMI, 615 N. Wolfe St., Baltimore, MD 21205.

0301-4622/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2004.02.007

defined tertiary interactions, apparently having disordered hydrophobic cores. It has been speculated that the reduction of sequence diversity results in the lack of specific interactions in the hydrophobic core necessary to form a unique native structure [12–14]. An oft cited analogy is that dominoes can be laid out in many different two dimensional grids, but the interactions in all the grids are identical. In contrast, a jigsaw puzzle has only one arrangement with optimal interactions between the pieces. If the idea behind this analogy is correct, the diversity in sequence in natural proteins may be important in making specific tertiary interactions and resulting in a single, relatively well-defined native state. Indeed, some designs with more varied cores do seem to have well-defined core conformations.

This issue has also been studied in natural proteins by reducing the complexity of the hydrophobic core sequence in Rop, T4 lysozyme and Cro repressor. A Rop mutant with the hydrophobic core containing only leucine folded into a molten globule state [15]. In contrast, Rop with methionines in the core appeared to fold normally and 2 T4 lysozyme mutants with 7 and 10 methionine substitutions in the hydrophobic core were found to have well-defined native structures [16]. Similarly, a mutant of Cro repressor, with 11 out of 13 core residues substituted by leucine, was able to form a structure like wild-type [17]. On the other hand, another group engineered a human acidic fibroblast growth factor triple mutant increased core symmetry, although it had more than one amino acid type in the core. This mutant folded to a well defined native state, but showed significant deviation from the two-state denaturation behavior of the wild-type protein [18]. The newly apparent folding intermediate in this mutant might be a stabilized molten globule, although this is not clear.

In order to understand the relationship between hydrophobic core sequence complexity and protein structure the effect of simplifying the hydrophobic core in other proteins need to be examined. Recent work in our laboratory on staphylococcal nuclease affords a unique opportunity to examine the effects of simplified cores relative to other possible packing mutations.

Isoleucine, leucine and valine are the three major hydrophobic residues conserved in the hydrophobic cores of staphylococcal nuclease and its homologs [19]. We present here the results of simplifying the core of staphylococcal nuclease by substituting six selected core residues with a single type of residue: isoleucine, leucine or valine. The six positions mutated here are fewer in number than some of the examples cited above but we recently published the stabilities of 12 single [20], 44 double [21], 64 triple and 32 quadruple mutants [22], representing all possible permutations of isoleucine, leucine and valine at 2 overlapping sets of 4 positions selected from these same 6 positions. In addition, quadruple and quintuple mutants which assess the stability effects of the consensus residues found in homologs of staphylococcal nuclease have also been constructed [19]. These mutants are the largest library of packing mutants ever thermodynamically characterized. This library allows us to firmly place the effects of core simplification in the larger context of other possible packing mutations at these same sites.

1. Materials and methods

The mutants were prepared by multiple cycles of Kunkel mutagenesis, and the mutated genes were transferred into an expression vector, expressed and purified as described previously [23]. Guanidine hydrochloride denaturation and data analysis was carried out as previously described [24].

Enzyme activities of the leucine core and related mutants were measured by the method of Cuatrecasas et al. [25] at room temperature of 24 °C. The activity of the wild-type protein was measured at the same time. The amount of protein was adjusted in the range of 0.13–0.26 μg/ml to obtain rates optimal for measurement.

The crystals of the leucine core and related mutants were obtained as described previously [23]. Data collection and structure refinement were also carried out as described previously [23]. Calculations of the root-mean-square deviations (RMSDs) of the backbones between experimental or calculated structures were performed in the CNS software package [26] and the C terminus, residues 1–6, and N terminus, 142–149, which are not visible in the electron density, were not used. In all RMSD calculations the 42–50 loop was also excluded because of poor electron density. Any residues with

a side chain substitution between two structures were excluded as well.

2. Results

The six core residues mutated in this study are valines 23, 66 and 99, isoleucines 72 and 92, and leucine 25. The triple mutant L25V/I72V/I92V has six valines at these positions. The all leucine core mutant contains five mutations, V23L/V66L/I72L/I92L/V99L. The all isoleucine core mutant is a quadruple mutant, V23I/L25I/V66I/V99I. This is summarized in Table 1. In the course of constructing these three mutants, eight previously unpublished mutants containing substitutions at the selected six core sites were also constructed and they were examined as controls.

2.1. Stability from GuHCl denaturation

The results of guanidine hydrochloride denaturation of these mutants are summarized in Table 2. All multiple mutants were less stable than wild-type. The stability changes of these multiple mutants relative to wild-type (ΔΔG) are compared to the sum of the stability effects from the corresponding single mutants ($\Sigma\Delta\Delta G_{singles}$). The interaction energy, $\Delta^nG_{int\text{-}1}$, discussed in detail elsewhere [21,22,27], is the difference between ΔΔG and $\Sigma\Delta\Delta G_{singles}$. The superscript n indicates the number of mutant side chains in the multiple mutant and the subscript 1 indicates that the interaction energy takes into account only the effects of single mutants and not, for example, the interactions between component triple mutants in a quadruple mutant. (In the case of double mutants, the subscript 1 is usually dropped for Δ^2G_{int} since there are no lower order component multiple mutants.) When $\Delta^nG_{int\text{-}1}$ is positive, there are favorable interactions among the substitutions and the multiple mutant is more stable than would be predicted from the effects of the single mutants. Similarly, a negative $\Delta^nG_{int\text{-}1}$ is presumably due to strains introduced by the mutations and indicates that the multiple mutant is less stable than expected from the effects of the component single mutations. It should be noted that interaction energies are empirical, not a fundamental thermodynamic property and that several assumptions are made in their calculation [28,29], but they nevertheless are powerful in their explanation of experimental observations of packing effects and protein stability [19].

Most of these multiple mutants had interaction energies near zero. In other words, the stability of the multiple mutant was well predicted by adding the effects of the component single mutants. However, three mutants did have larger magnitude interaction energies. The most notable of these was the all leucine core, V23L/V66L/I72L/I92L/V99L. It was 1.1 kcal/mol less stable than predicted by summing the stability effects of the constituent single mutants. The other two mutants with significant interaction energies were V23L/L25V/V66I/V99I and V23I/V66L/I72L/I92L/V99L, being 0.4 and 0.5 kcal/mol, respectively, less stable than predicted by the sum of the stability effects of their component single mutants.

Table 1
Hydrophobic core sequences in wild-type staphylococcal nuclease and simplified core mutants

	23	25	66	72	92	99
Wild-type	Val	Leu	Val	Ile	Ile	Val
Valine core mutant	–	Val	–	Val	Val	–
Leucine core mutant	Leu	–	Leu	Leu	Leu	Leu
Isoleucine core mutant	Ile	Ile	Ile	–	–	Ile

2.2. Crystal structures of the leucine core and related mutants

The crystal structures of the all leucine core, V23L/V66L/I72L/I92L/V99L, and a related mutant, V23I/V66L/I72L/I92L/V99L, were solved to resolution of 1.65 and 1.70 Å, respectively. The proteins are unliganded, with no inhibitor or calcium ion present. These structures have been deposited in the Protein DataBank under accession numbers 1IHZ (V23L/V66L/I72L/I92L/V99L) and 1II3 (V23I/V66L/I72L/I92L/V99L).

The overall structures of these two mutants are very similar to that of wild-type (PDB number 1EY0) [19]. The RMSDs of the backbones in the structures of the leucine core (1IHZ) and related mutant (1II3) relative to that of wild-type are 0.62 and 0.55 Å, respectively. The structural changes in the leucine core mutant (V23L/V66L/I72L/I92L/V99L) relative

Table 2
Solvent denaturation data of the simplified core mutants and the related mutants

Mutants	ΔG_{H_2O}[a]	C_m[b]	m_{GuHCL}[c]	$\Delta\Delta G$[d]	$\Sigma\Delta\Delta G_{single}$[e]	$\Delta^n G_{int\text{-}1}$[f]	$\Delta^n G_{int\text{-}2}$[g]	ΔCH_2[h]
L25V/I72V/I92V[i]	2.2	0.28	1.16	−3.2	−3.4	0.2	0.2	−3
L25I/I72V/I92V	2.0	0.25	1.19	−3.4	−3.3	−0.1	0.1	−2
V23L/V66L/I72L/I92L/V99L[i]	2.8	0.43	0.99	−2.6	−1.5	−1.1	n.d.	3
V23I/V66L/I72L/I92L/V99L	3.3	0.58	0.86	−2.1	−1.6	−0.5	n.d.	3
V66L/I72L/I92L/V99L[j]	4.1	0.65	0.98	−1.3	−1.4	+0.1	0.6	2
V23I/L25I/V66I/V99I[i]	2.3	0.42	0.84	−3.1	−3.1	0.0	−0.1	3
V23I/L25V/V66I/V99L	2.1	0.40	0.80	−3.3	−3.2	−0.1	0.0	2
V23L/L25I/V66I/V99L	2.2	0.39	0.87	−3.2	−3.0	−0.2	−0.3	3
V23L/L25V/V66I/V99L	1.9	0.32	0.90	−3.5	−3.1	−0.4	−0.5	2
V23I/L25I/V66I/V99L	2.1	0.38	0.85	−3.3	−3.2	−0.1	0.3	3
V23L/V66I/V99L	4.2	0.71	0.90	−1.2	−1.3	0.1	−0.3	3
V23I/V66I/V99L	3.8	0.77	0.76	−1.6	−1.4	−0.2	−0.8	3
Wild-type	5.4	0.82	1.00	–	–	–	–	–

[a] Free energy difference between native and denatured states in the absence of denaturant in units of kcal/mol. Error is estimated to be ±0.1 kcal/mol.

[b] Midpoint concentration (concentration of guanidine hydrochloride at which half of the protein is denatured) in units of M. Error is estimated to be ±0.01 M.

[c] Slope value (change in free energy with respect to change in guanidine hydrochloride concentration) expressed relative to wild-type value of 6.53 kcal/(mol·M). Error is estimated to be ±0.02.

[d] Difference in free energy between the free energy of the mutants and the free energy of wild-type. $\Delta\Delta G=\Delta G_{H_2O}$ (mutant)−5.4 (WT). Error is estimated to be ±0.2 kcal/mol.

[e] The sum of the $\Delta\Delta G_{single}$ values of the component single mutants [20].

[f] $\Delta^n G_{int\text{-}1}=\Delta\Delta G-\Sigma\Delta\Delta G_{single}$.

[g] $\Delta^n G_{int\text{-}2}$: when n=3, $\Delta^3 G_{int\text{-}2}=\Delta\Delta G-\Sigma\Delta\Delta G_{double}+\Sigma\Delta\Delta G_{single}$. When n=4, $\Delta^4 G_{int\text{-}2}=\Delta\Delta G-\Sigma\Delta\Delta G_{double}+2\Sigma\Delta\Delta G_{single}$. Lack of data for the close 23/92 double mutant pairs made calculation less reliable for some mutants, which are indicated with n.d. (not determined).

[h] Difference in the number of methylene groups between the mutant and wild-type.

[i] The all valine, all leucine and all isoleucine core mutants are shown in bold. Other similar multiple mutants are grouped below the most closely related homogenous core.

[j] The denaturation parameters for this mutant were previously published [22].

to wild-type are similar to the changes in the closely related mutant V23I/V66L/I72L/I92L/V99L.

The major hydrophobic core of staphylococcal nuclease is defined by a five strand anti-parallel β-barrel and two α-helixes. Residues of 23 and 25 are located on β-strand 2, and residues 72 and 92 are on β-strands 4 and 5, respectively. Residues 66 and 99 are located on helixes 1 and 2, respectively.

In these two structures, substitution of valine with leucine or isoleucine at positions 23, 66 and 99 increases the total side chain size by three methylene groups. This causes the backbone near these three residues to move outwards from the core. The backbone changes near residues 66 and 99, which are located in α-helixes, only affect the immediately neighboring residues. The backbone movement at position 23 is the largest, about 0.5 Å (Fig. 1). This movement extends to the whole β-strand 2, and also causes β-strands 1 and 3, hydrogen bonded to either side of strand 2, to move consequently. Substitutions of isoleucine with leucine at positions 72 and 92 do not change the side chain size. The backbone of residue 72 overlaps that of wild-type, but the backbone of residue 92 moves slightly outwards from the core.

Other interesting structural changes are observed at residues 19 and 21. The aspartates 19 and 21 are located in the turn that connects β-strands 1 and 2. The side chain orientation of aspartate 19 is completely changed and that of aspartate 21 is moved away from its original position in wild-type structure (Fig. 1). Both these aspartates are highly conserved and aspartate 21 is known to be one of the residues required to bind the calcium ion essential for nuclease activity [25,30]. In the presence of inhibitor and calcium ion, aspartates 19 and 21 adopt different side chain conformations from those in the wild-type structure without inhibitor and calcium ion. When

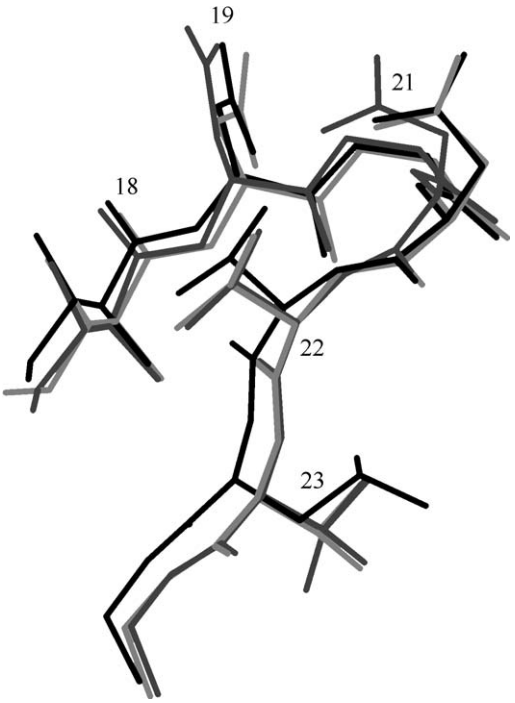

Fig. 1. Comparison of side chain conformations of residues 19 and 21 in the structures of unliganded wild-type staphylococcal nuclease (1EYO, dark gray), wild-type nuclease liganded with Ca^{2+} and inhibitor (1SNC, light gray) and the all leucine core mutant (1lHZ, black). The motion of the main chain at position 23 is the most pronounced such change in the mutant protein. Figure prepared with MolScript [39].

the structures of the leucine core and related mutant are compared to the wild-type structure with inhibitor and calcium ion (PDB number 1SNC) [30], it is found that the side chain conformations of aspartates 19 and 21 in the two mutants are very similar to those in 1SNC (Fig. 1).

2.3. Enzyme activity

The enzyme activities of V23L/V66L/I72L/I92L/V99L and V23I/V66L/I72L/I92L/V99L were the same as that of wild-type within experimental error.

3. Discussion

Three mutants, each with only one type of residue (valine, leucine or isoleucine) at six positions in the hydrophobic core of staphylococcal nuclease, have been constructed in our study. Other mutants which either substitute the same positions with different side chains, or represent a subset of the substitutions in the all valine, all leucine or all isoleucine mutants were also constructed. These mutants and the large library of mutants previously constructed at these positions [19–22] allow us to examine the effects of replacing these six side chains with a single amino acid type relative to other possible mutations that also alter packing in the hydrophobic core.

Not surprisingly, all the multiple mutants are less stable than wild-type, usually by very significant amounts (Table 2). The key point to note is that the stability of the all valine, all leucine and all isoleucine mutants is not particularly reduced compared to the other mutants in Table 2 nor do the values of C_m and m_{GuHCl} stand out as unusual in comparison.

While informative, examining only the stability losses experienced by these mutants does not exploit all the available information. Previously, we found that interaction energies correlated strongly with evolutionary fitness as assessed by the occurrence of particular side chains at these six positions in homologous proteins [19]. Favorable interaction energies for a pattern of side chain substitution was much better correlated with frequent appearance for that particular pattern in related proteins than the simple stability of the mutant with that pattern. We interpreted this to be due to the fact that the overall stability of the protein measures not only the interactions of the mutated side chains with each other but also the interactions of those residues with the rest of the protein. Interaction energies, to a first approximation, report only the effects of interactions between the mutated residues. Thus, regardless of whether the overall protein context is favorable for a given multiple mutation pattern, interaction energies report whether or not interactions between the mutated side chains are stabilizing or not.

In the mutants presented for the first time here interaction energies near zero, i.e. additive stability effects, are found in most cases. There are no particularly favorable or unfavorable interactions introduced among mutated side chains. However, three proteins do have unusually low interaction energies relative to the other mutants in Table 2. The all leucine mutant V23L/V66L/I72L/I92L/V99L and the closely related V23I/V66L/I72L/I92L/V99L mutant have

Δ^nG_{int-1} energies of −1.1 and −0.5 kcal/mol, respectively. They are significantly less stable than the sum of the stability effects calculated from the component single mutants. One other protein, V23L/L25V/V66I/V99I, has an interaction energy of −0.4 kcal/mol.

However, viewed in the larger context of the known range of interaction energies these values are not particularly unusual. This is illustrated graphically in Fig. 2 where we compare the Δ^nG_{int-1} values of the all leucine, all isoleucine and all valine core mutants to those of all other packing mutants in Table 2 as well as the approximately 150 multiple mutants previously examined at these six sites [19,21,22]. It is obvious that the near zero interaction energies for the all isoleucine and the all valine mutants are not at all unusual. The negative interaction energy of the all leucine mutant is notable but not unique.

This comparison is somewhat biased, since our earlier packing studies [21,22] have shown that stability effects are correlated with the change in methylene groups upon mutation (ΔCH_2). A loss of methylene units generally lead to a more positive value for the interaction energy and a gain of methylene units to a more negative value. However, in our previous work [19,21,22], the largest gain or loss of methylenes by any mutant was just two. The only mutants we have yet constructed where three methylene units were gained or lost are those presented in Table 2. To increase the numbers of mutants for comparison in Fig. 3A, we show the all valine mutant as well as all other mutants which have lost either two or three methylene units relative to wild-type. In Fig. 3B, the all leucine and all isoleucine mutants as well as all other mutants that have gained either two or three methylene units. The all valine and all isoleucine exhibit quite comparable interaction energies to packing mutants with similar changes in number of methylene units. The all leucine mutant is at the low end of the observed range of behavior, but is not a true outlier. We also considered the possibility that the high number of substitutions in the mutants presented here relative to those we have previously examined might be a factor. However, after either limiting comparison to mutants with similar numbers of sub-

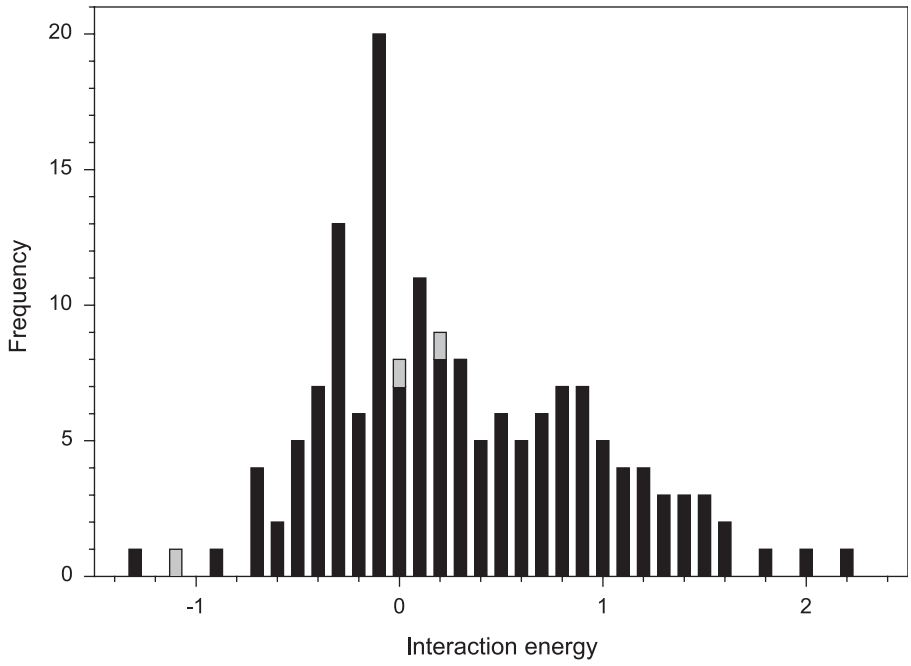

Fig. 2. Histogram of interaction energies (Δ^nG_{int-1}) for various packing mutants of staphylococcal nuclease. The number of times mutants are found with a given interaction energy is indicated by the frequency axis. All mutants with two or more substitutions at these six sites from previously published work [19,21,22] as well as all mutants in Table 2 are included. The all leucine, all valine and all isoleucine mutants from Table 2 are indicated in gray. All other packing mutants are indicated in black.

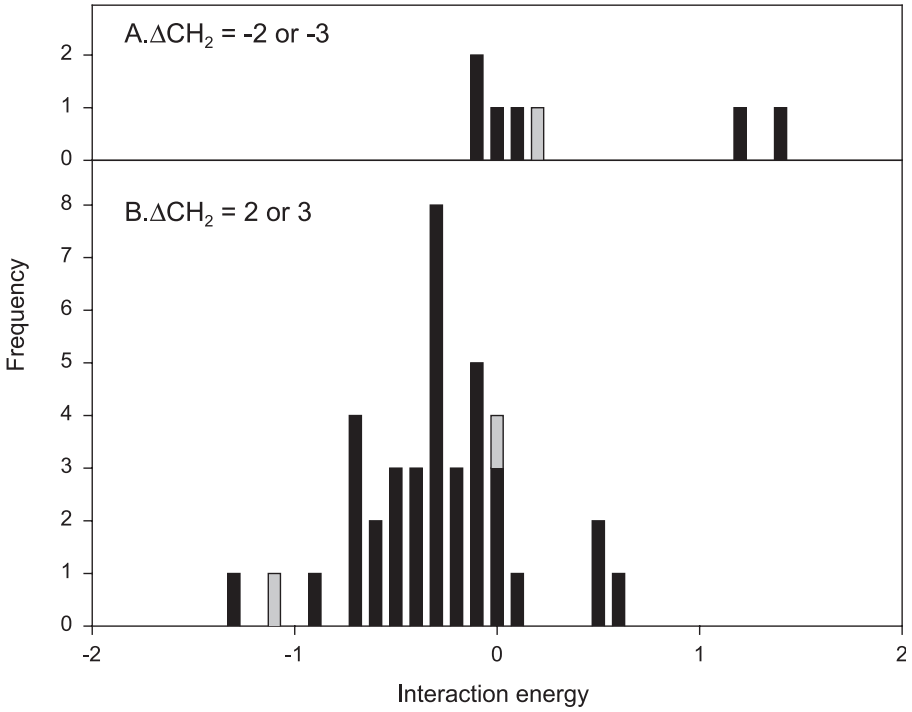

Fig. 3. Histogram of interaction energies ($\Delta^nG_{int\text{-}1}$) for packing mutants of staphycoccal nuclease sorted by the change in the number of methylene units with respect to wild-type. The all valine mutant is potted (gray) in panel A along with all other mutants which have lost two or three methylene units (black). The all leucine and all isoleucine mutants are plotted (gray) in panel B along with all other mutants which have gained two or three methylene units (black).

stitutions or after normalizing $\Delta^nG_{int\text{-}1}$ by dividing by *n* the number of substitutions, the all valine and all isoleucine mutants are, if anything, even less remarkable (data not shown). The all leucine is still at the low end of the observed range of behavior, but again we do not find it remarkably low.

Previously, we had shown that the effects of triple and quadruple mutants were very well predicted by taking into account not only the effects of the component single mutants, but also the interaction energies of the component double mutants (Δ^2G_{int}) [22]. Further, we have shown that the interaction energies of double mutants that are not in close proximity to one another (i.e. a $C_\beta - C_\beta$ distance of more than 8 Å) are near zero. We have constructed most, but not all component double mutants that make up these multiple mutants. Generally, the distance between the two residues is far enough that the assumption of a zero value of Δ^2G_{int} seems warranted for those pairs where we do not have actual data. The sole exception to this is residue pair 23/92, for which we do not have data, but which are close enough that non-zero interaction energies are likely. Accordingly, we calculated the residual interaction energies for each mutant after taking in account the effects of single mutants and double mutants ($\Delta^nG_{int\text{-}2}$ in Table 2) assuming zero values of Δ^2G_{int} for double mutants we have not constructed. Because of their proximity, no $\Delta^nG_{int\text{-}2}$ value is shown for multiple mutants which have mutations at both 23 and 92. Most values of $\Delta^nG_{int\text{-}2}$ are very near to zero, indicating that the effects of the single and double mutants well predict the stability of the multiple mutant, whether it is all one type of side chain or not.

In short, there is no evidence that cores limited to a single amino acid type at these six positions are particularly destabilized or energetically unusual merely because they are simplified.

This, however, does not rule out structural changes that are not energetically obvious. In a very

interesting series of studies performed in Richards' laboratory [31–33], packing effects in the major hydrophobic core of staphylococcal nuclease were studied by chemically creating different length unnatural amino acids at position 23. The valine at position 23 was first replaced by a cysteine. A series of unnatural straight and cyclic sulfides were connected to cysteine at position 23 by a disulfide bond. The crystal structures of these mutants showed that the nuclease hydrophobic core could accommodate a side chain as large as 1-*n*-propyl cysteine disulfide without perturbation of the surrounding residues. When the side chains were longer than 1-*n*-propyl cysteine disulfide, the packing around these side chains was still very similar to the wild-type. The over-packing was compensated for by the displacement of a helix in the global structure. However, in a very striking and unusual result, the unnatural side chain itself became disordered and no longer had recognizable electron density. This disordering of the side chain is akin to what is thought to happen to all side chains in a molten globule.

Could mutants with simplified cores fold into native-like structures but have a disordered cores like the unnatural amino acids placed at valine 23 or perhaps even be full fledged molten globules? The latter seems unlikely. Again, the guanidine hydrochloride denaturation behavior of these proteins was not unusual in any way compared the hundreds of staphylococcal nuclease mutants we have examined in the past. Second, all the proteins here that were stable enough to examine thermal denaturation behavior had high cooperativity (data not shown) similar to wild-type and other mutants, which is different from the typically lower cooperativity of thermal denaturation found in molten globules [34–38]. However, as the example of the disordered unnatural side chain at V23C plainly illustrates, the possibility of local disorder must be taken seriously.

The crystal structures of the two mutants with the lowest interaction energies, the all leucine mutant V23L/V66L/I72L/I92L/V99L and the closely related V23I/V66L/I72L/I92L/V99L mutant, were accordingly solved to high resolution. The overall structures of both mutants are similar to that of wild-type. Most importantly, the core side chains, mutated and unaltered alike, have defined conformations. The *B* values of the backbone and side chains for the core residues are all lower than the average *B* values of the backbone and side chains of the entire structure, with the exception of position 92 which has a side chain *B* value similar to the average. Therefore, the data of the crystal structures and thermal denaturation of the leucine core and related mutants indicate that the leucine core and related mutants have unique and defined structures in the hydrophobic core without even local disorder and certainly are not molten globules. The disorder in unnatural amino acids at position 23 observed by others [31–33] is perhaps an artifact of that particular system rather than being a general packing issue, possibly reflecting why such amino acids are not found in nature.

In conclusion, the present study demonstrates that staphylococcal nuclease with large numbers of isoleucine, leucine or valine in the hydrophobic core can still fold into well defined native state. No particular destabilization of the protein which could be attributed to sequence simplification was observed. The marked stability changes found in these mutations seem reasonably attributed to poor packing of the core. Crystal structures of the two mutants with the worst interaction energies had no special disorder of the side chains, either local or global. A single packing arrangement was still found in the core. This, and the similar results found in other proteins [15–17], indicate that the molten globule like behavior of some designed proteins can not be purely attributed to simplified cores with low sequence diversity. If the domino versus jigsaw puzzle analogy cited in the introduction is correct, it appears it is also because of symmetry at higher levels of structure.

Acknowledgements

Funding for this work was provided by NIH grant GM52714 and NCRR COBRE Grant 1 P20 RR15569.

References

[1] F.M. Richards, Protein stability: still an unsolved problem, Cell. Mol. Life Sci. 53 (1997) 790–802.

[2] M. Levitt, M. Gerstein, E. Huang, S. Subbiah, J. Tsai, Protein folding: the endgame, Ann. Rev. Biochem. 66 (1997) 549–579.
[3] J. Lee, K. Lee, S. Shin, Theoretical studies of the response of a protein structure to cavity-creating mutations, Biophys. J. 78 (2000) 1665–1671.
[4] J.G. Northey, A.A. Di Nardo, A.R. Davidson, Hydrophobic core packing in the SH3 domain folding transition state, Nat. Struct. Biol. 9 (2002) 126–130.
[5] Q. Wang, A.M. Buckle, A.R. Fersht, Stabilization of GroEL minichaperones by core and surface mutations, J. Mol. Biol. 298 (2000) 917–926.
[6] H.K. Schroder Leiros, N.P. Willassen, A.O. Smalas, Structural comparison of psychrophilic and mesophilic trypsins. Elucidating the molecular basis of cold-adaptation, Eur. J. Biochem. 267 (2000) 1039–1049.
[7] F.M. Richards, Areas, volumes, packing and protein structure, Annu. Rev. Biophys. Bioeng. 6 (1977) 151–176.
[8] R.B. Hill, D.P. Raleigh, A. Lombardi, W.F. DeGrado, De novo design of helical bundles as models for understanding protein folding and function, Acc. Chem. Res. 33 (2000) 745–754.
[9] S.F. Betz, W.F. DeGrado, Controlling topology and native-like behavior of de novo-designed peptides: design and characterization of antiparallel four-stranded coiled coils, Biochemistry 35 (1996) 6955–6962.
[10] G.A. Lazar, T.M. Handel, Hydrophobic core packing and protein design, Curr. Opin. Chem. Biol. 2 (1998) 675–679.
[11] J.R. Desjarlais, T.M. Handel, New strategies in protein design, Curr. Opin. Biotechnol. 6 (1995) 460–466.
[12] T.M. Handel, S.A. Williams, W.F. DeGrado, Metal ion-dependent modulation of the dynamics of a designed protein, Science 261 (1993) 879–885.
[13] C.J. Tsai, B. Ma, Y.Y. Sham, S. Kumar, R. Nussinov, Structured disorder and conformational selection, Proteins 44 (2001) 418–427.
[14] B.I. Dahiyat, S.L. Mayo, Probing the role of packing specificity in protein design, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 10172–10177.
[15] M. Munson, S. Balasubramanian, K.G. Fleming, A.D. Nagi, R. O'Brien, J.M. Sturtevant, L. Regan, What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties, Protein Sci. 5 (1996) 1584–1593.
[16] N.C. Gassner, W.A. Baase, B.W. Matthews, A test of the "jigsaw puzzle" model for protein folding by multiple methionine substitutions within the core of T4 lysozyme, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 12155–12158.
[17] J.R. Desjarlais, T.M. Handel, De novo design of the hydrophobic cores of proteins, Protein Sci. 4 (1995) 2006–2018.
[18] S.R. Brych, S.I. Blaber, T.M. Logan, M. Blaber, Structure and stability effects of mutations designed to increase the primary sequence symmetry within the core region of a beta-trefoil, Protein Sci. 10 (2001) 2587–2599.
[19] J. Chen, W.E. Stites, Packing is a key selection factor in the evolution of protein hydrophobic cores, Biochemistry 40 (2001) 15280–15289.
[20] J.B. Holder, A.F. Bennett, J. Chen, D.S. Spencer, M.P. Byrne, W.E. Stites, Energetics of side chain packing in staphylococcal nuclease assessed by exchange of valines, isoleucines, and leucines, Biochemistry 40 (2001) 13998–14003.
[21] J. Chen, W.E. Stites, Energetics of side chain packing in staphylococcal nuclease assessed by systematic double mutant cycles, Biochemistry 40 (2001) 14004–14011.
[22] J. Chen, W.E. Stites, Higher-order packing interactions in triple and quadruple mutants of staphylococcal nuclease, Biochemistry 40 (2001) 14012–14019.
[23] J. Chen, Z. Lu, J. Sakon, W.E. Stites, Increasing the thermostability of staphylococcal nuclease: implications for the origin of protein thermostability, J. Mol. Biol. 303 (2000) 125–130.
[24] J.M. Schwehm, W.E. Stites, Application of automated methods for determination of protein conformational stability, Methods Enzymol. 295 (1998) 150–170.
[25] P. Cuatrecasas, S. Fuchs, C.B. Anfinsen, Catalytic properties and specificity of the extracellular nuclease of *Staphylococcus aureus*, J. Biol. Chem. 242 (1967) 1541–1547.
[26] A.T. Brunger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Crystallography and NMR system: a new software suite for macromolecular structure determination, Acta Crystallogr., D Biol. Crystallogr. 54 (Pt 5) (1998) 905–921.
[27] A. Horovitz, A.R. Fersht, Strategy for analysing the co-operativity of intramolecular interactions in peptides and proteins, J. Mol. Biol. 214 (1990) 613–617.
[28] S. Boresch, M. Karplus, The meaning of component analysis: decomposition of the free energy in terms of specific interactions, J. Mol. Biol. 254 (1995) 801–807.
[29] A.E. Mark, W.F. van Gunsteren, Decomposition of the free energy of a system in terms of specific interactions. Implications for theoretical and experimental studies, J. Mol. Biol. 240 (1994) 167–176.
[30] P.J. Loll, E.E. Lattman, The crystal structure of the ternary complex of staphylococcal nuclease, Ca^{2+}, and the inhibitor pdTp, refined at 1.65 A, Proteins 5 (1989) 183–201.
[31] R. Wynn, C.L. Anderson, F.M. Richards, R.O. Fox, Interactions in nonnative and truncated forms of staphylococcal nuclease as indicated by mutational free energy changes, Protein Sci. 4 (1995) 1815–1823.
[32] R. Wynn, P.C. Harkins, F.M. Richards, R.O. Fox, Mobile unnatural amino acid side chains in the core of staphylococcal nuclease, Protein Sci. 5 (1996) 1026–1031.
[33] R. Wynn, P.C. Harkins, F.M. Richards, R.O. Fox, Comparison of straight chain and cyclic unnatural amino acids embedded in the core of staphylococcal nuclease, Protein Sci. 6 (1997) 1621–1626.
[34] O.B. Ptitsyn, Molten globule and protein folding, Adv. Protein Chem. 47 (1995) 83–229.
[35] I. Nishii, M. Kataoka, Y. Goto, Thermodynamic stability of the molten globule states of apomyoglobin, J. Mol. Biol. 250 (1995) 223–238.
[36] D. Xie, E. Freire, Molecular basis of cooperativity in protein folding: V. Thermodynamic and structural conditions for the stabilization of compact denatured states, Proteins 19 (1994) 291–301.

[37] W. Pfeil, V.E. Bychkova, O.B. Ptitsyn, Physical nature of the phase transition in globular proteins. Calorimetric study of human alpha-lactalbumin, FEBS Lett. 198 (1986) 287–291.
[38] D.A. Dolgikh, L.V. Abaturov, I.A. Bolotina, E.V. Brazhnikov, V.E. Bychkova, R.I. Gilmanshin, O. Lebedev Yu, G.V. Semisotnov, E.I. Tiktopulo, O.B. Ptitsyn, et al., Compact state of a protein molecule with pronounced small-scale mobility: bovine alpha-lactalbumin, Eur. Biophys. J. 13 (1985) 109–121.
[39] P.J. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, J. Appl. Crystallogr. 24 (1991) 946–950.