Accelerated Publications

Toward a Simplification of the Protein Folding Problem: A Stabilizing Polyalanine α-Helix Engineered in T4 Lysozyme†,‡

X-J. Zhang, W. A. Baase, and B. W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute, and Department of Physics, University of Oregon, Eugene, Oregon 97403

Received October 29, 1990; Revised Manuscript Received December 18, 1990

ABSTRACT: In an attempt to simplify the protein folding problem, and also to further investigate the role of alanine as a helix-stabilizing residue, a series of alanines was introduced within the α-helix that includes residues 126–134 of T4 lysozyme. In wild-type lysozyme this α-helix contains alanine residues at positions 129, 130, and 134. Mutant lysozymes with alanines substituted at positions 128, 131, 132, and 133, either as single substitutions or in selected combinations, were constructed by oligonucleotide-directed mutagenesis. With the exception of the replacement of Leu 133, which is buried within the hydrophobic core of the protein, all the variants were more stable than wild-type lysozyme. The variant with alanines substituted at positions 128, 131, and 132 (E128A/V131A/N132A), which incorporates the sequence Ala 128-Ala 129-Ala 130-Ala 131-Ala 132-Leu 133-Ala 134, has a melting temperature 3.3 °C above that of wild-type lysozyme. Determination of the crystal structure of this mutant lysozyme shows that the replacement of Glu 128, Val 131, and Asn 132 with alanine causes α-helix 126–134 to rotate 3.4° about an axis parallel to its own axis. This rotation seems to be triggered primarily by the loss of a hydrogen bond between Asn 132 and Ser 117 and is associated with the repacking of several side chains at the interface between α-helix 126–134 and the adjacent α-helix 115–122. Presumably because of these conformational adjustments, there is a synergistic effect such that the stabilization arising from the triple substitution E128A/V131A/N132A is slightly greater than the sum of the stabilization energies due to the constituent single plus double substitutions E128A and V131A/N132A. The results show that at least for one α-helix in T4 lysozyme the replacement of solvent-exposed residues with alanines provides a method to increase the stability of the protein. Also, the finding that a series of alanines can be introduced into T4 lysozyme shows that this might be a way to simplify the protein folding problem.

Although it has long been recognized that the amino acid sequence of a protein determines its three-dimensional structure (Epstein et al., 1963), recent work has made it clear that certain amino acids are more important than others in the folding process. At some positions, typically the solvent-exposed or mobile sites in the folded protein, amino acids can be interchanged almost at will with little apparent effect on folding or stability (Perutz & Lehmann, 1968; Hecht et al., 1983; Shortle & Lin, 1985; Reidhaar-Olson & Sauer, 1988; Alber et al., 1988; Bowie et al., 1990). These amino acids seem to be unimportant in protein folding. On the other hand, sites at which amino acid replacements substantially destabilize folded proteins appear to be mostly restricted to the buried or rigid parts (Alber et al., 1987), suggesting that the amino acids at these positions are important in determining the folded conformation.

The observation that many amino acids in a protein sequence can be freely replaced with little if any effect on protein stability leads one to consider the protein folding problem in a new light. What fraction of the amino acids in a given polypeptide sequence is, in fact, essential for the successful folding of the protein? Is it 75%? Is it 50%? Could it be as low as, say, 20%? To restate the problem in a different way, could one take a known protein and replace a large number of "nonessential" amino acids with alanine and yet still have a folded functional protein? Such a "polyalanine protein" would truncate all nonessential side chains and allow one to focus on those parts of the amino acid sequence that were critical for the folding process. Cunningham and Wells (1989) and collaborators have introduced a related approach, "alanine-scanning mutagenesis", to identify specific side chains that are important in hormone-receptor recognition.

There are several potential difficulties with the successful construction of a polyalanine protein. For example, it might be possible to replace residues one at a time with alanine, but a series of consecutive replacements might lead to nonnative conformations. Also, the net free energy of stabilization of folded proteins is very small, typically 5–20 kcal/mol (Creighton, 1983). Therefore, a series of alanine replacements, each of which was only slightly destabilizing, could, in combination, quickly exceed this small "margin of stability". Successive introduction of alanines might also reduce the solubility of the protein. An alternative approach, which might alleviate some of the above problems, would be to restrict replacements not just to alanine but to a small subset of the 20 natural amino acids.

In the present report, which can be considered as a modest first step toward a polyalanine protein, we explore the con-

† This work was supported in part by grants from the NIH (GM21967) and the Lucille P. Markey Charitable Trust.

‡ Coordinates of the mutant lysozyme have been deposited in the Brookhaven Protein Data Bank.

* Address correspondence to this author at the Institute of Molecular Biology.

0006-2960/91/0430-2012$02.50/0

Table I: Sequences of Mutant Lysozymes[a]

	α-helix 126–134									
lysozyme	126	127	128	129	130	131	132	133	134	135
wild type	Trp	Asp	Glu	Ala	Ala	Val	Asn	Leu	Ala	Lys
E128A			Ala	(Ala)	(Ala)				(Ala)	
V131A				(Ala)	(Ala)	Ala			(Ala)	
V131A/N132A				(Ala)	(Ala)	Ala	Ala		(Ala)	
L133A				(Ala)	(Ala)			Ala	(Ala)	
E128A/V131A/N132A			Ala	(Ala)	(Ala)	Ala	Ala		(Ala)	
solvent accessibility	0.27	1.10	0.76	0.00	0.14	0.78	0.34	0.01	0.37	0.85

[a] The table shows the locations of the alanines introduced in the different mutant lysozymes. Alanines in parentheses are present in both wild-type and mutant variants. The last line gives the fraction of the wild-type side chain accessible to solvent. The solvent accessibility was calculated (Lee & Richards, 1971) as the ratio of the side-chain area accessible to surface in wild-type lysozyme relative to the side-chain area accessible to solvent in an extended model peptide with the same amino acid sequence (Alber et al., 1987).

sequences of introducing a series of alanines within an α-helix of bacteriophage T4 lysozyme.

There is also a second motivation for such an experiment, namely, to further examine the role of the alanine residue in contributing to α-helix stability. Recent studies with short model peptides have shown that alanine has a high "helix-forming tendency" (Marqusee et al., 1989; Merutka & Stellwagen, 1990; Merutka et al., 1990; Padmanabhan et al., 1990; O'Neil & DeGrado, 1990; Lyu et al., 1990). Also the substitution Val 131 → Ala increased the stability of T4 lysozyme presumably because of the incorporation of the alanine within an α-helix (Daopin et al., 1990). Two questions addressed in the present study are (1) whether other replacements of the type Xaa → Ala within α-helices may provide a general method to increase protein stability and (2) whether the consequences of an alanine substitution are context dependent.

The results show not only that a mutant T4 lysozyme with a series of five consecutive alanines within an α-helix will fold normally but also that the protein is actually more stable than wild type.

Design of the Mutations

Because we were interested in the stabilizing effect of a series of consecutive alanines within an α-helix, we sought sites at which one or more alanines were already present. In wild-type T4 lysozyme there are three Ala-Ala pairs, all within α-helices. Ala 41-Ala 42 is within α-helix 39–50, Ala 73-Ala 74 is within α-helix 60–80, and Ala 129-Ala 130 is within α-helix 126–134. We chose the latter for two reasons. First, it was already known that the mutant Val 131 → Ala increased the stability of T4 lysozyme (Daopin et al., 1990). Second, the adjacent residues in the α-helix, Glu 128 and Asn 132, were solvent exposed. Therefore, by substituting Val 131, Glu 128, and Asn 132, up to five consecutive alanines could be obtained within the 126–134 α-helix without obviously disrupting interactions that exist in the wild-type lysozyme structure. Such replacements will be the principal subject of this report. Extending further outward toward the ends of the α-helix, Leu 133 is buried within the core of the protein and is followed by another alanine (Ala 134). Toward the amino terminus of the α-helix, Asp 127 is largely solvent exposed but is in a position to interact with the "α-helix dipole". Asp 127 may also interact with Arg 154. The first residue in the α-helix, Trp 126, is about 75% solvent inaccessible.

Experimental Procedures

Construction of the mutant lysozyme Val 131 → Ala (V131A) has been described previously (Dao-pin et al., 1990). Additional single mutants E128A and L133A were prepared in a similar manner by oligonucleotide-directed mutagenesis (Kunkel et al., 1987). The double mutants E128A/V131A and V131A/N132A as well as the triple mutant E128A/V131A/N132A were also constructed to yield the sequences shown in Table I. Procedures for DNA sequencing, cloning, and protein purification were as described (Alber & Matthews, 1987; Muchmore et al., 1989; Poteete et al., 1991). Each of the mutants gave a single peak when analyzed by reverse-phase high-pressure liquid chromatography.

Crystals of all of the mutants were obtained under conditions similar to those used for wild-type lysozyme (Weaver & Matthews, 1987). High-resolution data were collected by oscillation photography (Rossmann, 1979; Schmid et al., 1981) and the structures refined by using the TNT package (Tronrud et al., 1987).

Stabilities of the wild-type and mutant proteins were determined by monitoring the change in circular dichroism (CD) at 223 nm as a function of temperature as described by Dao-pin et al. (1990) except as follows: van't Hoff analyses of the CD melting curves were used to determine thermodynamic parameters with the nonlinear minimization program NONLIN, kindly provided by Prof. Michael Johnson (Johnson & Frasier, 1985), and a thermodynamic model where the heat capacity change, ΔC_p, is invariant with temperature.

The difference in melting temperatures, ΔT_m, for each protein variant was calculated relative to wild type. In the solvent used, namely, 0.20 M KCl adjusted to pH 2.0 by addition of HCl, the wild-type T_m was determined to be 40.77 ± 0.2 °C. This T_m differs slightly from the previously quoted values [cf. Nicholson et al. (1988), Karpusas et al. (1989), and Dao-pin et al. (1990)]. Since dT_m/dpH is at least +12 °C/pH unit at pH 2, an inaccuracy in the value of the pH remains a likely cause of part of this shift. This possibility is supported in that a similar shift is seen for the T_m of V131A in this solvent (Dao-pin et al., 1990).

Values of $\Delta\Delta G$, the relative change in free energies associated with a given mutation, were calculated by using the $\Delta T_m{}^*\Delta S$ approximation, which is accurate for small changes in T_m and ΔS (Becktel & Schellman, 1987). Correct ΔT_m values are important in this calculation while absolute values of T_m are not. To this end, melts were always done in the same cell, with the same solvent, as close together as possible in time and including the reference protein, in this case wild type, which was prepared in parallel with the mutants.

Results

Each of the mutants that includes the substitution Glu 128 → Ala has measured activity (Tsugita et al., 1968) about 70% that of wild-type lysozyme. It was previously shown that the replacement Glu 128 → Lys slightly destabilizes the protein (Schellman et al., 1981) and reduces activity to 4% of wild type, presumably due to interference with substrate binding (Grutter & Matthews, 1982). The other replacements do not

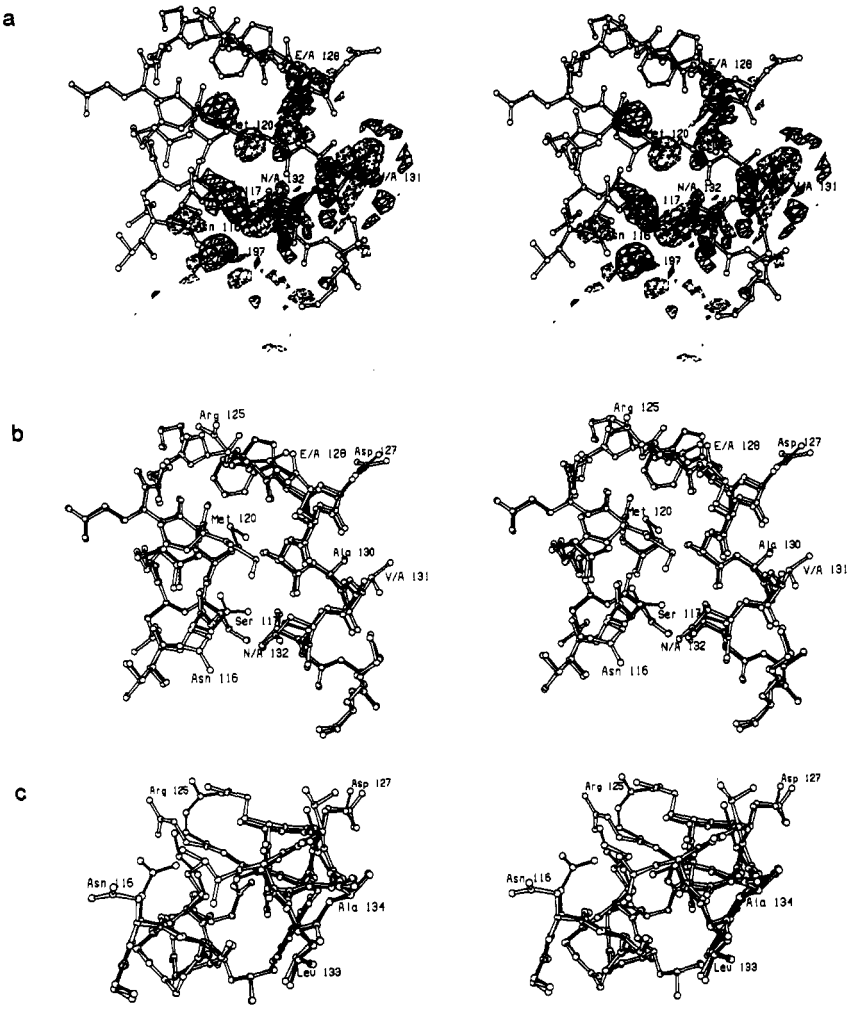

FIGURE 1: (a) Map showing the difference in electron density between the triple alanine mutant lysozyme E128A/V131A/N132A and wild-type lysozyme superimposed on the structure of wild-type lysozyme. Amplitudes are the differences between the observed structure amplitudes of mutant and wild-type lysozyme. Phases are from the refined structure of wild-type lysozyme. Resolution = 1.7 Å. Contours are drawn at $+3.5\sigma$ (solid lines) and -3.5σ (dashed lines), where σ is the average electron density throughout the unit cell. (b) Superposition of the refined structure of mutant lysozyme E128A/V131A/N132A (solid bonds) on wild-type lysozyme (open bonds). (c) Superposition of the refined structure of E128A/V131A/N132A (solid bonds) on wild-type lysozyme (open bonds), as in (b), but viewed so as to show the rotation of the 126–132 α-helix.

appear to have a significant effect on activity.

The results presented here will be restricted to a summary of the stabilities of the mutant lysozymes together with a description of the crystal structure of the triple alanine mutant. Table II summarizes the melting temperatures of the mutant lysozymes and the inferred changes in the free energy of unfolding relative to wild type. With the exception of the mutant with the internal Leu 133 → Ala replacement, all the variants unfold and refold reversibly and all are more stable than wild type. The Leu 133 → Ala replacement is much less stable than wild type, melting about 15 °C lower at pH 2.

The structure determination of the triple alanine replacement E128A/V131A/N132A was based on 13942 reflections to 1.7-Å resolution. Cell dimensions were $a = b = 61.3$ Å and $c = 96.3$ Å. The structure was refined to an R-value of 16.4% with a root-mean-square deviation of bond lengths and bond angles from their ideal values of 0.015 Å and 2.3°, respectively.

The map showing the difference electron density between the mutant structure and wild-type lysozyme is shown in Figure 1a. As expected, there are negative features due to the replacement of Glu 128, Val 131, and Asn 132 with alanine. In the case of Glu 128 there is little density at the extremity of the side chain, consistent with the fact that the thermal factors of the atoms in the carboxyl group have an average value of 77 Å^2, indicating high mobility. There are, however, additional features in Figure 1a that are clearly significant. First, there is a pair of positive and negative peaks indicating a repositioning in the side chain of Met 120. Second, there is positive density at the position occupied by solvent 197 in the native structure. A very similar density peak was also observed for the V131A/N132A double mutant (data not shown). It is presumed that the reduction in size of the residue 132 side chain allows a chloride ion to bind with increased occupancy at the amino terminus of the α-helix

Table II: Stabilities of Mutant Lysozymes[a]

mutant	ΔH (kcal/mol)	ΔT_m (°C)	$\Delta\Delta G$ (kcal/mol)
WT	86.0		
E128A	85.0	0.6 ± 0.25	0.16
V131A	88.7	1.0 ± 0.25	0.26
E128A/V131A	93.0	1.5 ± 0.25	0.41
V131A/N132A	82.0	2.3 ± 0.25	0.62
E128A/V121A/N132A	88.3	3.4 ± 0.22	0.94

[a] ΔH is the enthalpy of unfolding at the melting temperature, T_m. ΔT_m is the difference between the melting temperature of the mutant and that of wild-type lysozyme. $\Delta\Delta G$ is the difference between the free energy of unfolding of the mutant and wild-type proteins estimated from the relationship $\Delta\Delta G = \Delta T_m \Delta S$ (Becktel & Schellman, 1987). ΔS was taken as 274 ± 12 cal/(deg·mol). A positive $\Delta\Delta G$ indicates that the mutant lysozyme is more stable than wild type.

115–122 (Figure 1b). [Because this bound chloride ion also interacts with Lys 48 from a neighboring lysozyme molecule in the crystal lattice, it is not clear whether or not chloride is bound to a significant degree in solution [cf. Nicholson et al. (1988)].]

Finally, the difference map has features suggesting that there is movement of the 126–134 α-helix as a whole. The refinement of the mutant structure confirms that this is indeed the case. As can be seen in Figure 1c, α-helix 126–134 rotates about an axis essentially parallel to the helix axis, whereas the adjacent helix, 115–122, hardly moves at all.

Discussion

The most striking result of the present study is not so much that the protein tolerates the alanine replacements but that all but one of the four alanine substitutions actually increase the stability of T4 lysozyme. The exception is Leu 133 → Ala, which creates a large cavity within the hydrophobic core of the molecule and is very destabilizing. The structure of L133A and other replacements at residue 133 will be described elsewhere.

The two single replacements Val 131 → Ala (Daopin et al., 1990) and Glu 128 → Ala each increase the stability of the folded protein by 0.16–0.26 kcal/mol (Table II). We did not construct Asn 132 → Ala as a single mutant, but when combined with V131A, this replacement increased stability by 0.36 kcal/mol (Table II).

The structure of the triple mutant E128A/V131A/N132A reveals several structural changes above and beyond the removal of the three replaced side chains. The main trigger for these changes seems to be the Asn 132 → Ala replacement since the structural changes seen in the triple mutant are also observed in the double mutant V131A/N132A (data not shown) but are not seen (or are much reduced) in the known structures that do not include this substitution, V131A (Daopin et al., 1990) and E128A/V131A (data not shown). In wild-type lysozyme Asn 132 is hydrogen bonded to Ser 117 (distance 2.5 Å from 132 OD1 to 117 OD) (Figure 1b). Replacement of Asn 132 causes a loss of this hydrogen bond and causes the 126–134 α-helix to "roll" through an angle of 3.4° toward the vacated space. In association with the rotation of α-helix, there are substantial conformational adjustments (~2–4 Å) in the side chains of Asn 116, Met 120, and Arg 125 (Figure 1b). These three residues are at the interface of the 126–134 and 115–122 α-helices and are also partly exposed to solvent. Notwithstanding the loss of the hydrogen bond between Asn 132 and Ser 117, the replacement of Asn 132 with an alanine still increases the stability of the protein.

In the wild-type structure, Glu 128 may interact favorably with Arg 125 and perhaps also with the "helix dipole" of α-helix 115–122. Again, notwithstanding the loss of these possibly favorable interactions, replacement of Glu 128 with alanine still increases the net stability of the mutant protein relative to wild type. The stabilization arising from the Val 131 → Ala replacement is presumed to be due to the removal of close contacts between the valyl side chain and the backbone atoms within the adjacent turn of the α-helix (Daopin et al., 1990).

Additional studies of different amino acid replacements in different sequence and structural contexts, using both model peptides and intact proteins, will be necessary to fully understand all the factors that contribute to the "α-helix propensity" of a given amino acid. The present results do, however, support the idea that alanine is a strongly helix-forming residue (Marqusee et al., 1989; Daopin et al., 1990; Padmanabhan et al., 1990; O'Neil & DeGrado, 1990; Lyu et al., 1990). A survey of the amino acid sequences of 67 refined structures in the protein data bank (Bernstein et al., 1977) revealed that among 1050 alanines in a total of about 12000 residues there were 13 Ala-Ala-Ala triplets and no tetraalanine sequences. Of the 13 trialanines, 10 were within α-helices. For alanines in general, about 52% occur within α-helices (Chou & Fasman, 1974).

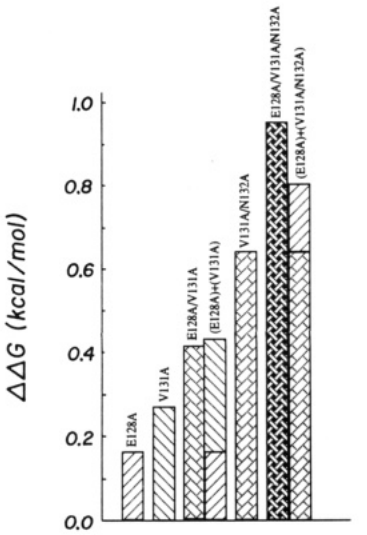

FIGURE 2: Histogram showing the stabilities of alanine-containing mutants relative to wild-type lysozyme. $\Delta\Delta G$ is the difference between the free energy of unfolding of the mutant lysozyme relative to that of wild-type lysozyme (Table II). Where available the free energies of the single and double mutants that constitute the multiple mutants are also included.

The present results do not distinguish between the enhanced helix stabilization of alanine relative to other amino acids, being due primarily to an effect on the folded or the unfolded state of the protein. It has been suggested that Val, Ile, and Thr are not favored in α-helices because their β-branched side chains tend to result in unfavorable steric interactions with the atoms in the preceding turn of the helix. At the same time, however, hydrophobic and entropic factors also contribute (Dao-pin et al., 1990). In the case of non-β-branched amino acids, their incorporation within an α-helix will, to some degree, restrict rotation about the C^{α}–C^{β} bond (i.e., χ_1), even if the residue is solvent exposed. Unless the energy of interaction of the side chain with the remainder of the α-helix compensates for the entropy cost of restricting the motion of the side chain, there will be a net reduction in stability. We assume that this must, in fact, be the normal situation, as it would serve to explain why alanine is the most stabilizing amino acid within an α-helix.

One of the intriguing aspects of the present study is the observation that the stabilization arising from the individual mutations is not always additive. In the case of the double mutant E128A/V131A, the increase in stability appears to be equal to the sum of E128A plus V131A. The stabilization arising from the triple mutant E128A/V131A/N132A, however, seems to be slightly greater than the sum of the constituent double and single mutants E128A plus V131A/N132A (Figure 2). For substituents that do not interact, the changes in energy are expected to be additive. Nonadditivity indicates that there is some form of interaction between the different sites (Carter et al., 1984). It would be surprising if the replacement of Val 131 caused nonadditive effects since the side chain of this residue is fully exposed to solvent (Figure 1b) and the Val 131 → Ala replacement causes very little change in the conformation of the protein (Dao-pin et al.,

1990). Indeed, the stabilization due to the E128A/V131A double mutant is equal to the sum of E128A plus V131A. In the triplet mutant, however, there are significant conformational changes (Figure 1b) that, as noted, seem to be mainly due to the Asn 132 → Ala replacement. We presume that the conformational adjustments associated with the replacement of Asn 132 with alanine (Figure 1b) create a somewhat more favorable environment for the Glu 128 → Ala replacement than exists in the wild-type lysozyme structure.

In formal terms one can relate the free energy changes associated with two single mutants, X and Y, to that of the double mutant (X/Y) by

$$\Delta\Delta G_{X/Y} = \Delta\Delta G_X + \Delta\Delta G_Y + \Delta G_{I,X/Y} \quad (1)$$

where $\Delta G_{I,X/Y}$ is the coupling energy between replacements at sites X and Y (Carter et al., 1984).

Thus, the coupling energy between sites 128 and 131 can be obtained as

$$\Delta G_{I,E128A/V131A} = \Delta\Delta G_{E128A/V131A} - \Delta\Delta G_{E128A} - \Delta\Delta G_{V131A} = -0.01 \text{ kcal/mol} \quad (2)$$

which is negligibly small.

For a triple mutant with coupling between all pairs of sites, one can write

$$\Delta\Delta G_{X/Y/Z} = \Delta\Delta G_X + \Delta\Delta G_Y + \Delta\Delta G_Z + \Delta G_{I,X/Y} + \Delta G_{I,Y/Z} + \Delta G_{I,X/Z} \quad (3)$$

If we assume that the replacement at site 131 is, in fact, not coupled to either site 128 or site 132 (i.e., $\Delta G_{I,E128A/V131A} = \Delta G_{I,V131A/N132A} = 0$), then the coupling between sites 128 and 132 can be obtained as

$$\Delta G_{I,E128A/N132A} = \Delta\Delta G_{E128A/V131A/N132A} - \Delta\Delta G_{E128A} - \Delta\Delta G_{V131A/N132A} = 0.16 \text{ kcal/mol} \quad (4)$$

With the same formalism, the stabilization expected from the single replacement N132A, which we have not yet constructed, can be predicted (assuming $\Delta G_{I,V131A/N132A} = 0$) as

$$\Delta\Delta G_{N132A} = \Delta\Delta G_{V131A/N132A} - \Delta\Delta G_{V131A} = 0.36 \text{ kcal/mol} \quad (5)$$

In terms of the original goal of simplifying the protein folding problem, we have shown that a series of residues in T4 lysozyme can be replaced with alanines with little effect on folding or catalytic activity. In the most extreme case to date, it is shown that the lysozyme that includes the penta-alanine sequence Ala 128-Ala 129-Ala 130-Ala 131-Ala 132 not only folds normally but is significantly more stable than wild type. In a helix-containing protein such as T4 lysozyme, it raises the possibility that alanine replacements of solvent-exposed residues within other α-helices might provide a general method to increase the net stability of the protein. In contrast, replacement of the buried residue, Leu 133, with alanine substantially destabilizes the protein, suggesting that this residue is important for folding.

Subsequent to the initial submission of this paper, it has been shown that replacements of three of the solvent-exposed residues in α-helix 40–49 of T4 lysozyme with alanines increase thermostability by amounts comparable to those reported here (D. W. Heinz, W. A. Baase, and B. W. Matthews, unpublished results).

ACKNOWLEDGMENTS

We thank Dr. John Schellman for helpful comments, Dr. Sun Dao-pin for advice on mutagenesis, and Joan Wozniak, Sheila Pepiot, and Joel Lindstrom for excellent technical assistance.

REFERENCES

Alber, T., & Matthews, B. W. (1987) *Methods Enzymol. 154*, 511–533.

Alber, T., Daopin, S., Nye, J. A., Muchmore, D. C., & Matthews, B. W. (1987) *Biochemistry 26*, 3754–3758.

Alber, T., Bell, J. A., Daopin, S., Nicholson, H., Wozniak, J. A., Cook, S., & Matthews, B. W. (1988) *Science 239*, 631–635.

Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1859–1877.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A., & Sauer, R. T. (1990) *Science 247*, 1306–1310.

Carter, P. J., Winter, G., Wilkinson, A. J., & Fersht, A. R. (1984) *Cell 38*, 835–840.

Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry 13*, 211–222.

Creighton, T. E. (1983) *Proteins: Structures and Molecular Properties*, Freeman, New York.

Cunningham, B. C., & Wells, J. A. (1989) *Science 244*, 1081–1084.

Daopin, S., Baase, W. A., & Matthews, B. W. (1990) *Proteins: Struct. Funct. Genet. 7*, 198–204.

Epstein, C. J., Goldberger, R. F., & Anfinsen, C. B. (1963) *Cold Spring Harbor Symp. Quant. Biol. 28*, 439–449.

Fermi, G., & Perutz, M. F. (1981) *Hemoglobin and Myoglobin*, Clarendon, Oxford.

Grutter, M. G., & Matthews, B. W. (1982) *J. Mol. Biol. 154*, 525–535.

Hecht, M. H., Nelson, H. C. M., & Sauer, R. T. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2676–2680.

Johnson, M. L., & Frasier, S. G. (1985) *Methods Enzymol. 117*, 301–342.

Karpusas, M., Baase, W. A., & Matthews, B. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8237–8241.

Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367–382.

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.

Lyu, P. C., Liff, M. I., Marky, L. A., & Kallenbach, N. R. (1990) *Science 250*, 669–673.

Marqusee, S., Robins, V. H., & Baldwin, R. L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5286–5290.

Matthews, B. W. (1987) *Biochemistry 26*, 6885–6888.

McGregor, M. J., Islam, S. A., & Sternberg, M. J. E. (1987) *J. Mol. Biol. 198*, 775–791.

Merutka, G., & Stellwagen, E. (1990) *Biochemistry 29*, 894–898.

Merutka, G., Lipton, W., Shalongo, W., Park, S.-H., & Stellwagen, E. (1990) *Biochemistry 29*, 7511–7515.

Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., & Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44–73.

Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature 338*, 651–656.

O'Neil, K. T., & DeGrado, W. F. (1990) *Science 250*, 646–651.

Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., & Baldwin, R. L. (1990) *Nature 344*, 268–270.

Perutz, M. F., & Lehmann, H. (1968) *Nature 219*, 902–909.

Poteete, A. R., Daopin, S., Nicholson, H., & Matthews, B. W. (1991) *Biochemistry* (in press).
Reidhaar-Olson, J. F., & Sauer, R. T. (1988) *Science 241*, 53–57.
Rossmann, M. G. (1979) *J. Appl. Crystallogr. 12*, 225–239.
Schellman, J. A., Lindorfer, M., Hawkes, R., & Grütter, M. G. (1981) *Biopolymers 20*, 1989–1999.
Schmid, M. F., Weaver, L. H., Holmes, M. A., Grutter, M. G., Ohlendorf, D. H., Reynolds, R. A., Remington, S. J., & Matthews, B. W. (1981) *Acta Crystallogr. A37*, 701–710.
Shortle, D., & Lin, B. (1985) *Genetics 110*, 539–555.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A34*, 489–501.
Tsugita, A., Inouye, M., Terzaghi, E., & Streisinger, G. (1968) *J. Biol. Chem. 243*, 392–397.
Weaver, L. H., & Matthews, B. W. (1987) *J. Mol. Biol. 193*, 189–199.

Identification of Vancomycin Resistance Protein VanA as a D-Alanine:D-Alanine Ligase of Altered Substrate Specificity†

Timothy D. H. Bugg,‡ Sylvie Dutka-Malen,§ Michel Arthur,§ Patrice Courvalin,§ and Christopher T. Walsh*,‡
Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, and Unitē des Agents Antibactēriens, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France
Received December 5, 1990

ABSTRACT: High-level glycopeptide resistance in *Enterococcus faecium* BM4147 is mediated by a 38-kDa protein VanA, whose amino acid sequence is related to Gram-negative D-alanine:D-alanine (D-Ala-D-Ala) ligases [Dutka-Malen, S., Molinas, C., Arthur, M., & Courvalin, P. (1990) *Mol. Gen. Genet. 224*, 364–372]. We report purification of VanA and demonstrate that it has D-Ala-D-Ala ligase activity but has substantially modified substrate specificity, compared with Gram-negative D-Ala-D-Ala ligases. VanA preferentially condenses D-Ala with D-Met or D-Phe, raising the possibility that its cellular role is to synthesize a modified cell-wall component, which is subsequently not recognized by vancomycin.

The vancomycin group of glycopeptide antibiotics, used clinically to treat life-threatening Gram-positive bacterial infections, functions by complexation of peptidyl-D-Ala-D-Ala termini on the surface of growing bacterial cell walls (Barna & Williams, 1984; Reynolds, 1989). Despite its clinical use for over 30 years, resistance to vancomycin has only recently emerged in strains of *Enterococcus faecium* and *Enterococcus faecalis* (Courvalin, 1990). Glycopeptide resistance is inducible by vancomycin and is associated with high-level expression of a 38–40-kDa protein that is associated with the cytoplasmic membrane (Nicas et al., 1989; Shlaes et al., 1989; Williamson et al., 1989). Subcloning of plasmid pIP816, conferring high-level vancomycin resistance in *E. faecium* strain BM4147 (Leclerq et al., 1988), allowed identification of the *vanA* gene that encodes one such inducible resistance protein designated VanA (Dutka-Malen et al., 1990). Sequencing of the *vanA* gene revealed 28–36% amino acid sequence homology with sequences of D-Ala-D-Ala ligase (Dutka-Malen et al., 1990), an enzyme that catalyzes synthesis of the D-Ala-D-Ala dipeptide for peptidoglycan assembly (Walsh, 1989).

Preliminary studies on the mechanism of resistance have excluded detoxification and impermeability of glycopeptides as possible mechanisms [reviewed in Courvalin (1990)]. In the light of its sequence similarity with Gram-negative D-Ala-D-Ala ligases, the following possible glycopeptide resistance mechanisms involving VanA were anticipated (Dutka-Malen et al., 1990): (i) overproduction of an *N*-acyl-D-Ala-D-Ala derivative, shown in vitro to compete with surface peptidyl-D-Ala-D-Ala residues for vancomycin binding (Nieto et al., 1972); (ii) complex formation with the peptidyl-D-Ala-D-Ala target, preventing binding of vancomycin; (iii) D,D-carboxypeptidase activity cleaving the terminal D-alanine residue, which is essential for binding of glycopeptides; (iv) biosynthesis of a modified peptidyl-D-Ala-D-Ala target from a D-Ala-D-Ala ligase of altered specificity. As an initial step toward elucidation of the molecular basis for vancomycin resistance, we describe purification of VanA, kinetic characterization and comparison with Gram-negative D-Ala-D-Ala ligases, and mechanistic studies to address the anticipated mechanisms.

Materials and Methods

Materials. The following materials were obtained from Sigma Chemical Co.: tris(hydroxymethyl)aminomethane base (Tris), ethylenediaminetetraacetic acid (EDTA), dithiothreitol (DTT), ampicillin, DL amino acids, D-alanine, D-methionine, D-phenylalanine, D-Ala-D-Ala, D-cycloserine, vancomycin, 6-aminopenicilloic acid, DL-Ala-DL-Phe, and L-Phe-L-Ala. The following materials were obtained from Boehringer Mannheim Biochemicals: phosphoenolpyruvate (PEP), pyruvate kinase/lactate dehydrogenase (PK/LDH), reduced nicotinamide adenine dinucleotide (NADH), ATP, and isopropyl thiogalactoside (IPTG). 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was purchased from U.S. Biochemicals. [U-^{14}C]-D-Alanine (40 mCi/mmol) was purchased from Amersham. Other chemicals and solvents were of reagent grade. Samples of UDPMurNAc-tripeptide and UDPMurNAc-pentapeptide were provided by Dr. J. van Heijenoort (University of Paris). *Salmonella typhimurium* D-Ala-D-Ala ligase A (Knox et al., 1989) and *Escherichia coli* D-Ala-D-Ala ligase B (Zawadzke et al., 1991) were prepared as previously described. *E. coli* D-Ala-D-Ala adding enzyme

† Supported in part by NSF Grant DMB 8917290 and a SERC/NATO Postdoctoral Fellowship (T.D.H.B.).
* Author to whom correspondence should be addressed.
‡ Harvard Medical School.
§ Institut Pasteur.

0006-2960/91/0430-2017$02.50/0