CRYSTAL PACKING

Acta Cryst. (1994). D**50**, 650-656

Packing Selection of *Bacillus lentus* Subtilisin and a Site-Specific Variant

BY JUDY L. DAUBERMAN,* GRANT GANSHAW, CURRAN SIMPSON, THOMAS P. GRAYCAR, SCOTT MCGINNIS AND RICHARD BOTT*

Department of Research and Development, Genencor International, Inc., 180 Kimball Way, South San Francisco, CA 94080, USA

(*Received* 3 *December* 1993; *accepted* 31 *January* 1994)

Abstract

The crystallization of a variant of *Bacillus lentus* subtilisin and the native enzyme was achieved using identical conditions. The variant *B. lentus* was found to grow in two crystal forms, form 1 and form 2, whereas the native *B. lentus* subtilisin enzyme crystallized in only one, form 1. Form 2 crystals, once obtained, were found to grow much more rapidly than form 1 crystals. The lattice contacts and structural changes giving both crystal forms have been examined. The results show that crystal form 2 has a more complex network of interactions. There is also a small surface conformational change in the form 2 structure relative to the native and variant form 1 crystals and at least two solvent molecules bound to the enzyme in crystal form 1 are displaced in crystal form 2. In addition, a site specific substitution in the variant at position 27 induces a 'short' lattice contact which does not exist in the native *B. lentus* or the form 2 variant *B. lentus*. These results suggest that in some circumstances engineered variants could be designed to crystallize more rapidly than the native enzyme.

Introduction

Subtilisin has been extensively studied in a number of laboratories (Wells & Estell, 1988) because it has commercial interest as a detergent additive and has served as a model system for protein engineering. Subtilisin from *Bacillus amyloliquefaciens* (subtilisin BPN′), *Bacillus licheniformis* (subtilisin Carlsberg) and *Bacillus lentus* have been crystallized and the high-resolution crystallographic structures have been determined many times. A number of crystal forms have been obtained for both subtilisin BPN′ and *B. lentus* subtilisin. In most of the previous cases the different crystal forms were obtained under quite different conditions or when different site-specific variants were crystallized.

We have obtained crystals of a variant of *B. lentus* subtilisin in two forms under identical conditions. Interestingly, we find that the native *B. lentus* subtilisin selectively grows in only one of these forms. The variant differs from the native enzyme at five positions. Four of these positions were the result of site-specific substitutions: (*a*) position 27, where arginine replaces lysine; (*b*) position 102, where tyrosine replaces valine; (*c*) position 121, where serine replaces asparagine; and (*d*) position 268, where alanine replaces threonine. In addition, the native enzyme is found to contain either serine or asparagine at position 85. Our native enzyme has asparagine and the variant has serine at position 85 (a consecutive numbering system from *B. lentus* subtilisin is used here instead of the usual alignment to the subtilisin BPN′ numbering).

The variant has twice the activity of the native enzyme in the presence of commercial detergents. The three-dimensional structures of the native and variant enzymes have been reported (Betzel *et al.*, 1992; Bott *et al.*, 1992; Goddette *et al.*, 1992; Teplyakov *et al.*, 1992). Table 1 compares the crystal forms obtained in these studies. One significant difference seen in the variant enzyme is a conformational change at the active site involving the segment of residues 125-129 which forms part of the substrate binding site. Previous research (Bott *et al.*, 1992) has shown this region to be selectively destabilized relative to the native enzyme. The variant has lower overall thermal stability relative to the native enzyme and it is believed that this destabilization is a factor in its improved performance. This performance advantage is perhaps as a consequence of increasing the flexibility at the active site.

In order to increase the understanding of the differences between the native and variant enzymes we have examined the factors which favor the appearance of two crystal forms for the variant under conditions in which the native enzyme is found to crystallize in only one, form 1.

Materials and methods

The crystals were grown by the hanging-drop method at 285 K. The purified proteins were buffer exchanged into 50 m*M* sodium acetate, pH 5.9, with 10 m*M* calcium chloride. The reservoir contained 29% saturated ammonium sulfate. The protein concentration in the

* To whom correspondence should be addressed.

Table 1. *Crystal forms of variant and native Bacillus lentus (alcalophilus) subtilisin*

	Betzel *et al.*	Teplyakov *et al.*			Goddette *et al.*	Bott *et al.*	
Space group	$P2_1$	$P2_12_12_1$			$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a (Å)	40.47	47.7	40.2	53.1	53.09	53.00	48.15
b (Å)	64.24	62.3	62.2	61.5	61.49	61.25	54.10
c (Å)	42.85	75.7	92.1	74.9	75.61	75.10	84.35
β (°)	118.8						
Precipitant	PEG 4000	PEG 6000			PEG 3350	29% sat. AS	
pH	6.5	6.0			5.7-5.9	5.9	

crystallization drop was approximately 6 mg ml^{-1}. Lower concentrations of protein tended to give fewer seeds and larger crystals. The method of seeding to obtain initial crystals was fiber-seeding which is the transfer of minute non-visible crystal seeds to a fresh protein drop.

The coordinates used were from refined models of native and variant *B. lentus* subtilisin (Bott *et al.*, 1992). Retrofitting the coordinates of a native or variant subtilisin determined in one crystalline form into a second crystalline form was accomplished by alignment of the homologous enzyme. In this instance, every 25th C^{α} atom of the retrofitted enzyme was aligned with the corresponding atoms of the reference molecule using the *RIGI* subroutine in *FRODO* (Jones, 1978) in the version written for the Evans and Sutherland 390. Lattice contacts were determined using the contact subroutine in *FRODO*.

Table 2. *Crystallographic and refinement statistics*

	Protein		
	Subtilisin *B. lentus* S85N	Subtilisin *B. lentus* K27R V102Y N121S T268A Form 1	Subtilisin *B. lentus* K27R V102Y N121S T268A Form 2
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell			
a (Å)	53.30	53.00	48.15
b (Å)	61.50	61.25	54.10
c (Å)	75.10	75.10	84.35
Resolution (Å)	10–1.8	10 1.7	10 2.0
No. of reflections	16480	19999	12204
Crystallographic *R* factor	0.148	0.148	0.154
R.m.s. deviation from ideality			
Bond length (Å)	0.014	0.015	0.014
Bond angle ( )	2.8	2.6	3.1

Results

The first crystals of *B. lentus* variant grew in space group $P2_12_12_1$ and had unit-cell dimensions *a* = 53.00, *b* = 61.20, *c* = 75.30 Å. These crystals have a block-like morphology and will be referred to as form 1. The second crystal form, form 2, spontaneously grew from a drop containing form 1 crystals. This second crystal form has a square-bipyramidal morphology, its space group is again $P2_12_12_1$ and has unit-cell dimensions *a* = 48.20, *b* = 54.15, *c* = 84.40 Å (Table 2). We have obtained native enzyme crystals in form 1 exclusively.

Fiber-seeding both the variant *B. lentus* and native *B. lentus* with crystal seeds from both forms 1 and 2 gave the result of variant *B. lentus* crystals of either form, and the native *B. lentus* crystals in form 1 only. The growth rate for form 2 crystals was fast relative to form 1. Form 1 crystals reached the 0.25 mm smallest diameter size in approximately 2–4 weeks using successive macro-seeding steps, whereas form 2 crystals reached the 0.25 mm size in 2–7 d using the same approach. Form 1 seeds develop more satellite crystals growing from either the fiber-seed or macro-seed, while form 2 crystals tend to grow as single crystals from both the fiber-seed and macro-seed crystal.

Lattice contacts

Form 1 crystals are stabilized by five pairs of lattice interactions listed in Table 3 for native and variant enzymes, respectively (unique interactions are represented in each table). Each molecule in form 1 crystals interacts with molecules in ten neighboring unit cells. There are 90 and 112 'good' van der Waals contacts for native and variant enzymes, respectively, along with 20 'short' contacts (defined as non-ionic contacts closer than 3.5 Å) for both native and variant. There are an additional 10 and 14 hydrogen bonds for variant and native enzymes, respectively, stabilizing form 1 crystals. Among the 'short' contacts none are closer than 3.10 Å in the native crystal or 3.03 Å in the variant.

Among the contacts listed in Table 3 only one region incorporates contacts involving a substituted residue

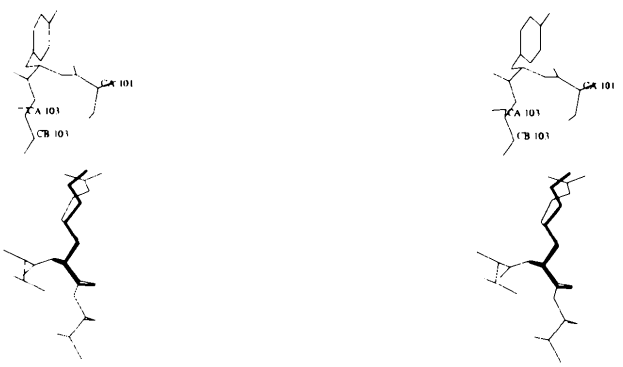

Fig. 1. Stereoview of the short contact in region 2 in crystal form 1 involving residue 27. The lysine side chain for the native enzyme is shown in thick lines.

Table 3. *Native (form 1) and variant (form 1) lattice interactions*

Region	Segments		Sym-metry	Unit cell	Interactions Good VDW	Short VDW	Good H	Short H
(*a*) Native (form 1)								
1	10	42	(i)	0 −1 −1	1	1	0	0
2	27	103	(ii)	−1 −1 0	1	0	0	0
3	51	116	(iii)	0 −1 0	7	0	0	0
		142–144						
4	39 42	176–179	(iv)	0 0 −1	18	4	2	0
	73–85	249 251						
5	129–134	198–210	(ii)	0 −1 0	18	5	3	0
	161 165							
		Total unique interactions			45	10	5	0
(*b*) Variant (form 1)								
1	10	42	(iv)	0 −1 −1	1	1	0	0
2	27	101–103	(iii)	−1 −1 0	1	1	0	0
3	51 52	116	(iii)	0 −1 0	9	0	0	0
		143–144						
4	39 42	176 179	(iv)	0 0 −1	18	4	3	0
	73 77	249 251						
5	129 135	198 210	(ii)	0 −1 0	27	4	4	0
	161 165							
		Total unique interactions			56	10	7	0

Symmetry codes: (i) x, y, z; (ii) $\frac{1}{2} - x, -y, \frac{1}{2} + z$; (iii) $\frac{1}{2} + x, \frac{1}{2} - y, -z$; (iv) $-x, \frac{1}{2} + y, \frac{1}{2} - z$.

Table 4. *Variant (form 2) lattice interactions*

Region	Segments		Sym-metry	Unit cell	Interactions Good VDW	Short VDW	Good H	Short H
1	19 24	127	(ii)	0 −1 −1	4	0	0	0
		156						
2	27	39	(vi)	1 0 −1	4	0	0	0
		77						
3	52	250–252	(i)	1 0 0	3	0	0	0
4	59	176	(iii)	0 0 0	4	0	0	0
	97	197–198						
5	75–77	111–115	(iv)	1 −1 −0	10	3	2	0
6	110	188 190	(iii)	0 1 0	9	1	2	0
	131–138	254–259						
7	175–182	203–205	(iii)	−1 0 0	6	6	1	1
		Total unique interactions			40	10	5	1

Symmetry codes: (i) x, y, z; (ii) $\frac{1}{2} - x, -y, \frac{1}{2} + z$; (iii) $\frac{1}{2} + x, \frac{1}{2} - y, -z$; (iv) $-x, \frac{1}{2} + y, \frac{1}{2} - z$.

between the native and variant enzyme. This substitution of arginine for lysine at position 27 in Table 3(*b*) introduces one short van der Waals contact for the variant enzyme in form 1 that is not present for the native enzyme in form 1 (Fig. 1). Variant form 2 crystals are stabilized by seven pairs of lattice interactions involving 14 unit cells listed in Table 4. Surprisingly, these crystals are stabilized by fewer van der Waals contacts (80 *versus* 90 and 112), the same number of 'short' van der Waals, 10 hydrogen bonds and one pair of 'short' hydrogen bonds (a 'short' hydrogen bond is defined as closer than 2.6 Å).

This comparison yields the striking result that five regions, albeit in different combinations, are involved in the lattice contacts for both crystal forms. These regions include the following residues 27, 51-52, 73-77, 129-135 and 175-179. Prominent among these segments are external loops such as 175-179 which will be discussed below. Composites of all the lattice contact regions involved in form 1 and form 2 crystals are presented in Table 5.

Table 5. *Comparison of variant, form 1 (left) with variant, form 2 (right) regional lattice contacts*

Region	Segments		Region	Segments	
1	10	42	1	19–24	127
2	27	101–103			156
3	51–52	116	2	27	39
		143–144			77
4	39–42	176–179	3	52	250–252
	73–77	249–251	4	59	176
5	129–135	198–210		97	197–198
	161–165		5	75–77	111–115
			6	110	188–190
				131–138	154–159
			7	175–182	203–205

Table 6. *Retrofitting coordinates*

Region	Segments		Sym-metry	Unit cell			Interactions Good VDW	Short VDW	Good H	Short H
(*a*) Variant (form 2) in variant crystal (form 1) lattice										
1	10–12	37 42	(iv)	0	−1	−1	2	2	1	0
2	27	103	(iii)	−1	−1	0	3	0	0	0
3	51 52	143–144	(iii)	0	−1	0	4	1	0	0
4	39	177 179	(iv)	0	0	−1	7	5	1	0
	75–77	249–251								
5	129–134	198 210	(ii)	0	−1	0	11	6	2	1
	161 164									
		Total unique interactions					27	14	4	1
(*b*) Variant (form 1) in variant crystal (form 2) lattice										
1	18–24	127 128	(ii)	0	1	−1	7	1	1	0
		156								
2	27	77	(iv)	1	0	−1	1	1	0	0
3	52–54	249 250	(i)	1	0	0	2	1	0	0
4	59	197 198	(iii)	0	0	0	2	5	1	0
	96 97									
5	75–77	111 115	(iv)	1	−1	−1	9	7	1	1
6	103	166								
	110	187 188	(iii)	0	1	0	9	13	1	2
	131–135	245								
		254 255								
7	177–182	203 209	(iii)	−1	0	0	7	10	1	1
		Total unique interactions					37	38	5	4
(*c*) Native (form 1) in variant crystal (form 2) lattice										
1	19 24	127	(ii)	0	1	−1	5	1	1	0
		156								
2	27	77	(iv)	1	0	−1	1	0	0	0
3	51–54	249 250	(i)	1	0	0	2	1	0	0
4	59	197	(iii)	0	0	0	0	1	0	0
5	75 77	111–115	(iv)	1	−1	−1	9	6	2	0
6	103	166								
	110	187–188	(iii)	0	1	0	8	11	0	2
	135	254 258								
7	177–182	203 205	(iii)	−1	0	0	6	11	0	2
		Total unique interactions					31	31	2	4

Symmetry codes: (i) x, y, z; (ii) $\frac{1}{2} - x, -y, \frac{1}{2} + z$; (iii) $\frac{1}{2} + x, \frac{1}{2} - y, -z$; (iv) $-x, \frac{1}{2} + y, \frac{1}{2} - z$.

Retro lattice contacts

To understand what happens to the variant enzyme when going from one form to another, and why the native enzyme does not appear in form 2, we have retrofitted coordinates of the variant enzyme determined in form 2 into the form 1 crystal lattice and *vice versa*. We have also positioned the coordinates of the native enzyme into the form 2 crystal lattice. The lattice contacts encountered when this was done are presented in Table 6. When either the variant determined in form 1 or the native enzyme is aligned onto the coordinates of the variant form 2, the number of 'short' van der Waals contacts increases dramatically to become equal or greater than the number of 'good' van der Waals contacts formed in the transformation.

Retrofitting the coordinates of the variant form 2 into the variant form 1 crystal lattice does not show this same effect. This is illustrated in Fig. 2 where the shift in region 175–182 from the form 2 determined structure results in a broadening of the gap between neighboring molecules relative to the interaction found for form 1 structures. Though the number of 'good' van der Waals contacts decrease, the number of 'short' van der Waals contacts does not appreciably increase. Many of these contacts can be seen to be alleviated by the subtle side-chain conformation differences seen between the form 1 determined structure and the form 2 determined structure (Fig. 3).

A comparison of the interactions made by the variant enzyme structure determined in the form 2 crystal *versus* those made for the variant structure determined in form 1, but then positioned in the form 2 crystal lattice, shows that a large number of short contacts are accumulated involving regions 4, 5, 6 and 7. Most of these are accommodated by rearrangements in side-chain conformations as can be seen by comparing the interaction in region 6 (Fig. 4).

Displacement of bound solvent in different crystal forms

In regions 6 and 7 we have found another significant difference between the structures of the variant enzyme determined in form 1 and form 2. Fig. 5 compares the

difference in residues 175-181 in variant form 1 and form 2 crystal structures. In the variant form 1 crystal structure one well ordered solvent molecule bridges the carbonyl O atoms of residues 190 and 254 on neighboring segments. In the form 2 crystal structure this molecule is displaced by the side chain of Gln135 with the NE side-chain atom now bridging the carbonyl O atoms of residues 190 and 254 on a neighboring molecule. A second ordered water may also be displaced by the Gln side chain as seen in Fig. 5. A similar phenomenon is seen when we examine the differences in residues 175-181 between the form 1 and form 2 crystal structures shown in Fig. 6. In the crystal structure of the form 1 variant a solvent stabilizes a loop structure by hydrogen bonding to the side chains of residues 175 and 177. In the form 2 variant crystal the OH of Tyr203 from a neighboring molecule displaces this ordered solvent (Fig. 7).

Thus, in the form 2 crystal packing 2–3 solvent atoms are displaced with the side chains of neighboring molecules forming the same interactions. There are no similar displacements when form 2 variant coordinates are positioned in the form 1 crystal lattice creating a net difference of 2–3 ordered solvent molecules between form 1 and form 2 crystals. This would be expected to introduce an entropic component favoring form 2.

Discussion

We have found that the variant crystallizes in two crystal forms which have very different packing arrangements. The second form differs from the first not only in fewer van der Waals interactions, but also that these interactions are distributed over 13 different neighboring unit cells (one pair of interactions occur with the reference unit cell), while the first form has five pairs of interactions linking ten neighboring unit cells. It is not clear whether the increased number of unit-cell linkages favor form 2 crystals. It might explain why crystal form 1 was encountered first, as this form involves fewer unit-cell linkages, hence, easier nucleation.

We know from an earlier study that the variant enzyme displays increased flexibility as evidenced by the disorder of the segment formed by residues 125-128 at the active site. This increased flexibility may be distributed throughout the molecule and may enable the conformational change that is seen between form 1 and form 2 variant structures for residues 175-181. Without this change, there would be numerous short van der Waals contacts occurring, as in our retrofitting of form 1 variant in the crystal form 2 lattice (Table 6*b*). Another probable factor is that the substitution of arginine for lysine at position 27 disfavors the variant crystallizing in form 1. Thus, the selection of the native

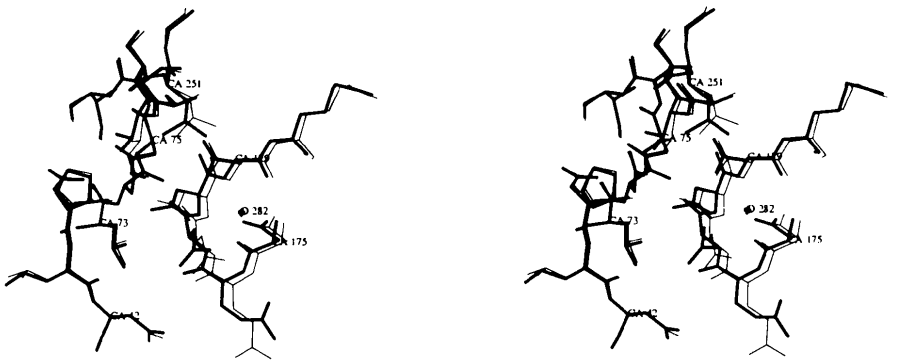

Fig. 2. Stereoview of crystal form 1 region 4 contacts. This view compares the coordinates of the variant enzyme determined in form 1 (thick lines) with the variant coordinates determined in crystal form 2 retrofitted into the form 1 crystal lattice (thin lines).

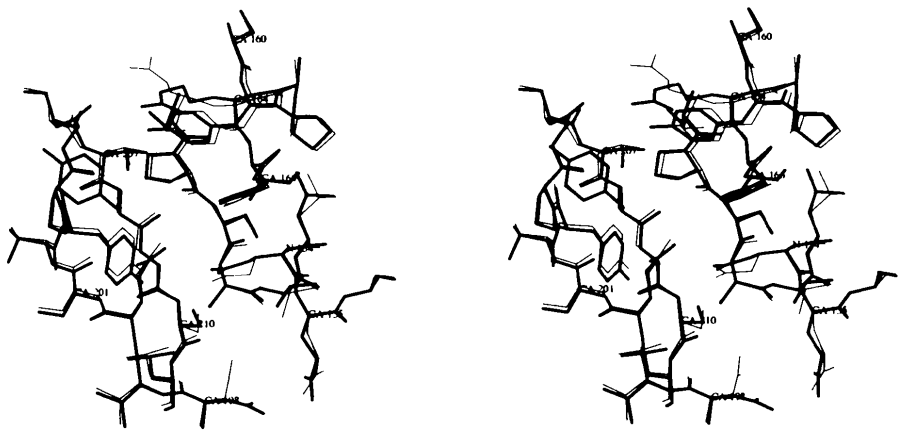

Fig. 3. Stereoview of the comparison of the variant determined in form 1 (thick lines) with the variant determined in form 2 (thin lines) for the residues involved in region 5, form 1 contacts.

crystallizing in form 1 only may be a result of two forces: (*a*) the presence of increased rigidity relative to the variant disfavoring the conformational change in residues 175-181, and (*b*) the absence of the additional short contact at position 27 that may nudge the variant to another crystal form. Conversely, the presence of arginine 27 and increased flexibility of the variant may facilitate the conformational change at positions 175-181.

Concluding remarks

We cannot account for the apparent preference of form 2 by the variant enzyme of the basis of the simple raw total number of good and bad van der Waals contacts (admittedly the limits were selected somewhat arbitrarily) and the number of good and bad hydrogen bonds. It is taken for granted that the short contacts reflect an artifact of the attempt to fit a single molecule to an electron density map reflecting an 'averaged molecule'. In reality, it is expected that short contacts represent close interactions which are accommodated by subtle and varied accommodations in lattice contacts throughout the crystal. Form 2 may be favored by an entropic component arising from the displacement of ordered solvent molecules seen in form 1.

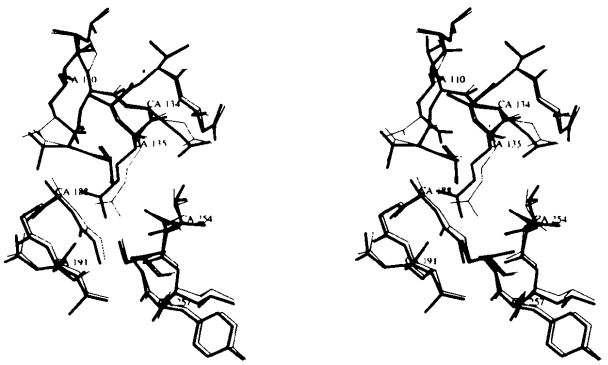

Fig. 4. Stereoview of the contacts in region 6 of form 2 crystals. The variant determined in form 2 is shown in thick lines with the retrofitted variant determined in form 1 shown in thin lines.

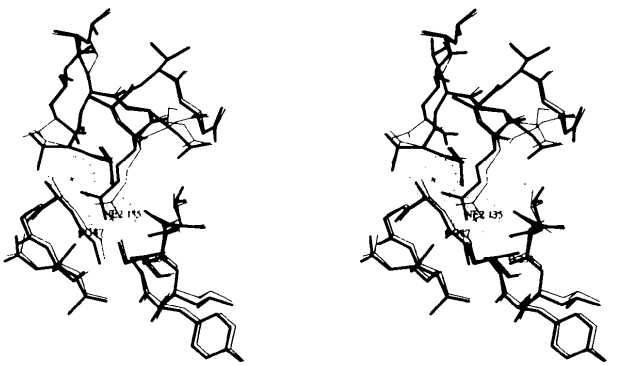

Fig. 5. Stereoview of region 6, form 2, with the ordered solvents found near residues 190 and 254 aligned with retrofitted form 1 variant coordinates. It can be seen that the side chain of Gln135 from a neighboring molecule in form 2 (thick lines) would displace one or both of these solvent molecules. This interaction is not a result of any conformational change between Gln135 (thin lines) in form 1 and form 2.

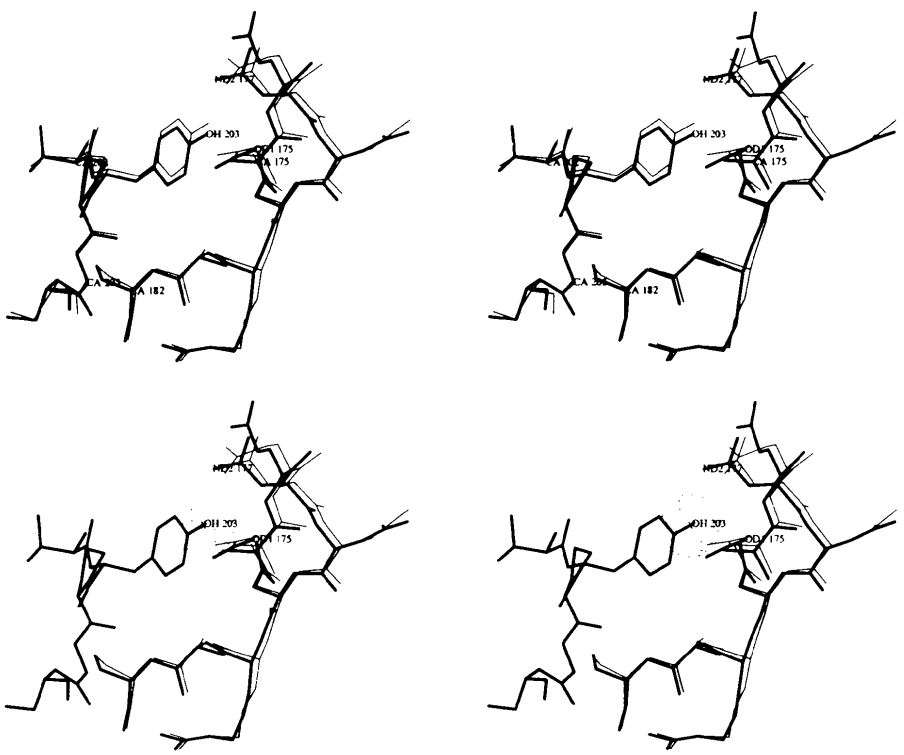

Fig. 6. Stereoview of region 7, form 2, with the variant determined in form 2 shown in thick lines and the variant determined in form 1 in thin lines. The shift in residues 175-188 represents a statistically significant shift between these residues in the two variant structures (Bott *et al.*, 1992).

Fig. 7. The position of an ordered solvent seen bound between residues 175 and 177 in form 1 (thin lines), and in dots the van der Waals surface. Relative to the corresponding residues of the variant determined in form 2 at region 7 (thick lines), Tyr203 of a neighboring molecule can be seen to displace this solvent in form 2.

It may be possible to exploit this propensity to adopt crystal forms which grow more rapidly for developing fast and economic recovery processes involving batch crystallization of recombinant DNA material.

References

BETZEL, C., KLUPSCH, S., PAPENDORF, G., HASTRUP, S., BRANNER, S. & WILSON, K. S. (1992). *J. Mol. Biol.* **223**, 427-445.

BOTT, R., DAUBERMAN, J., CALDWELL, R., MITCHINSON, C., WILSON, L., SCHMIDT, B., SIMPSON, C., POWER, S., LAD, P., SAGAR, H., GRAYCAR, T. & ESTELL, D. (1992). *Ann. N. Y. Acad. Sci.* **672**, 10-19.

GODDETTE, D. W., PAECH, C., YANG, S. S., MIELENZ, J. R., BYSTROFF, C., WILKE, M. E. & FLETTERICK, R. J. (1992). *J. Mol. Biol.* **228**, 580-595.

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268-272.

TEPLYAKOV, A. V., VAN DER LAAN, J. M., LAMMERS, A. A., KELDERS, H., KALK, K. H., MISSET, O., MULLENERS, L. J. S. M. & DIJKSTRA, B. W. (1992). *Protein Eng.* **5**, 413-420.

WELLS, J. A. & ESTELL, D. A. (1988). *Trends Biochem. Sci.* **13**, 291-297.