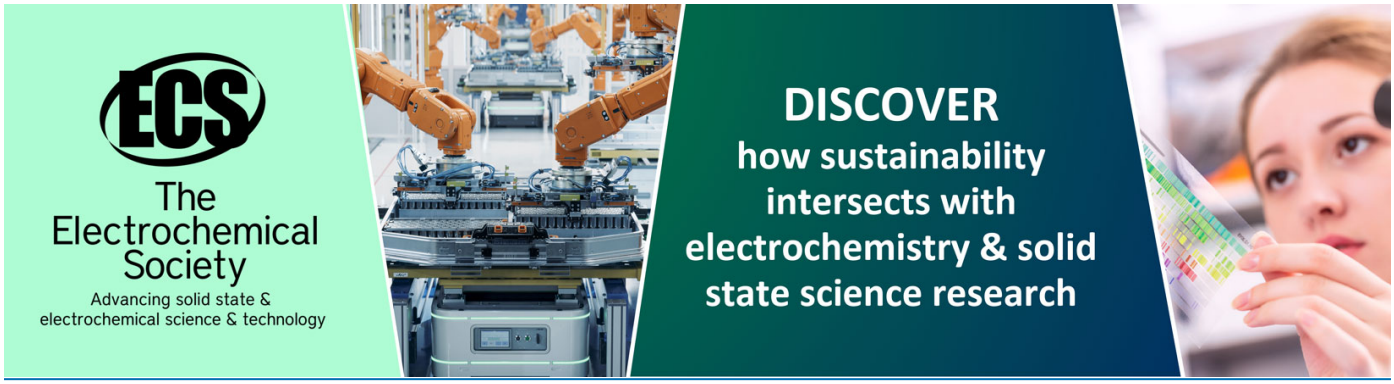
# Epitaxial protein inclusion in sinapic acid crystals





J. Phys. D: Appl. Phys. 26 (1993) 442–447. Printed in the UK

# Epitaxial protein inclusion in sinapic acid crystals

**Ronald C Beavis†§ and John N Bridson‡**

† Department of Physics and ‡ Department of Chemistry, The Memorial University of Newfoundland, St John's, Newfoundland, Canada A1B 3X7

Received 25 June 1992, in final form 11 November 1992

**Abstract.** We report on the inclusion of protein molecules in growing crystals of *trans*-3,5-dimethoxy-4-hydroxy-cinnamic acid (*trans*-sinapic acid). This material is important as a 'matrix' compound in matrix-assisted laser desorption, a promising new ion source for heavy molecule mass spectrometry. The crystal structure of *trans*-sinapic acid was determined. It belongs to the space group $P2_1/n$ and has a structure in which the planar *trans*-sinapic acid molecules are hydrogen bonded into extended sheets. Dyes were used to follow protein incorporation into growing *trans*-sinapic crystals. The staining pattern obtained demonstrated that protein molecules were included only on crystal faces parallel to these extended sheets. This result is the first demonstration that proteins can be incorporated into a growing crystal by a selective interaction with one of the crystal faces. The structure of the crystal plane that interacts with the protein was consistent with a 'hydrophobic' bonding of the protein to the crystal face.

## 1. Introduction

Evidence from protein ablation experiments suggests that proteins can be intimately mixed with some organic molecules, simply by adding them to an aqueous solution and drying [1–5]. These composite deposits produce isolated, monomolecular protein ions when illuminated with a pulsed ultraviolet laser in vacuum. The type of organic molecule ('matrix') is critical: only a very few matrix compounds produce protein ions. The ability of some compounds to assist protein ablation has been a puzzle, since the discovery of the matrix-assisted laser desorption effect for proteins. The matrix compound originally discovered (nicotinic acid) had several serious analytical drawbacks, chiefly its inability to tolerate physiological salt concentrations. The discovery of a group of *trans*-cinnamic acid derivatives that are suitable matrices for proteins (e.g. sinapic acid) has allowed the examination of mixtures of proteins by mass spectrometry without extensive purification because of their high tolerance to common contaminants [6]. Gentisic acid [5] (2,5-dihydroxybenzoic acid) and vanillic acid [7] (4-hydroxy-3-methoxybenzoic acid) also have similar properties.

The specificity of these materials for proteins in the face of high concentrations of contaminants suggests that these materials have a particular affinity for protein incorporation. Examination of the solid deposits formed by drying protein/matrix solutions shows that they consist of small crystals characteristic of the matrix compound. This observation implies that matrix crystals must incorporate protein molecules into their structure [5]. The defects produced by the protein in the matrix crystals must be at least the same size as the protein molecule. Defects of this size are most likely inclusions in the matrix crystal, rather than a true solid solution of protein and matrix molecules.

It is not generally appreciated that some of the best matrix materials had been previously identified as having a strong affinity for proteins. In studies of 'difficult to remove' contaminants in protein extracts, hydroxylated cinnamic acids [8] (in particular *para*-coumaric, caffeic and ferulic acids) and gentisic and vanillic acids [9] were

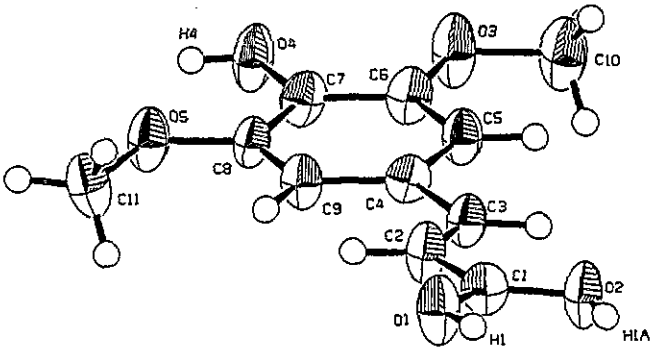



**Figure 1.** A drawing of a *trans*-sinapic acid molecule in the conformation adopted in its crystalline state, with the atoms labelled as in table 1.

§ To whom all correspondence should be addressed.

0022-3727/93/030442+06$07.50 

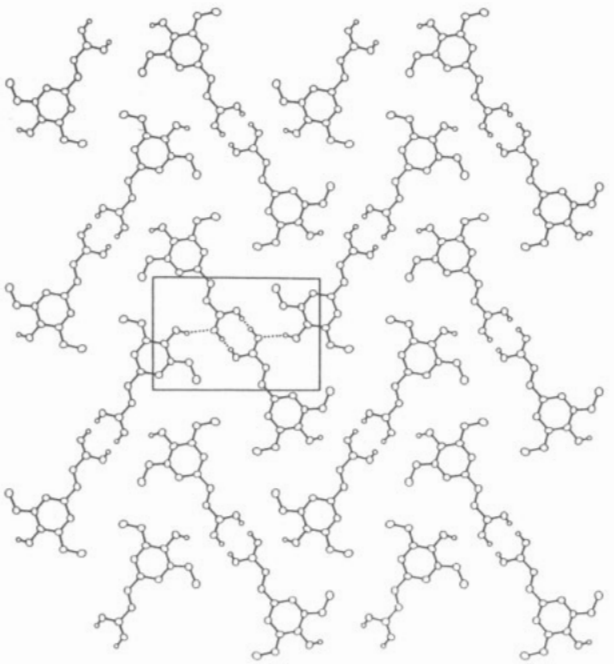


**Figure 2.** A view of the *trans*-sinapic acid crystal structure, perpendicular to its $(10\bar{3})$ plane, drawn by ORTEP II, using the experimentally determined crystal parameters.

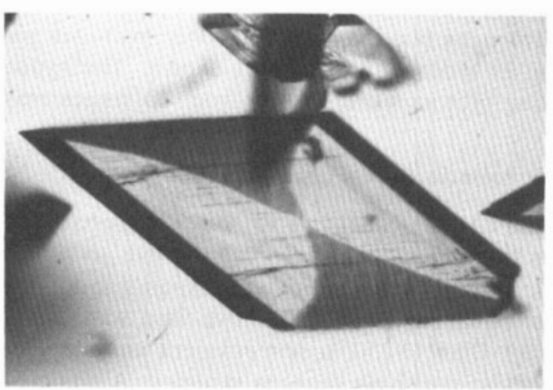


**Figure 3.** Crystals of *trans*-sinapic acid grown from a mother liquor containing protein and Coomassie Brilliant Blue G-250, demonstrating inhomogeneous protein inclusion (see section 2.2). The $(10\bar{3})$ face of the large crystal is at the base of the darkly stained (i.e. protein containing) hourglass zone and the (010) face is in the plane of the photomicrograph. The full width of the photomicrograph shown corresponds to a real distance of 0.9 mm.

shown to bind strongly to protein fractions obtained from plant leaves.

*Trans*-sinapic acid was chosen to study the mechanism for protein incorporation into matrix crystals because it is currently the most commonly used material for the analytical matrix-assisted laser desorption of proteins. Its crystal structure had not been determined and little was known about its most common crystal habits. It is also a matrix material with a high affinity for proteins in the face of high concentrations of other molecules in solution. It was hoped that determination of its crystal structure and comparison of that structure with those

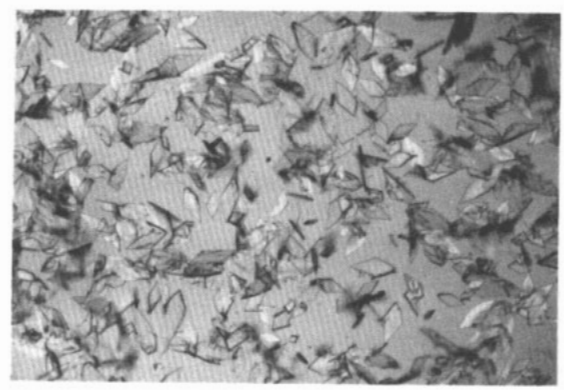


**Figure 4.** Crystals of *trans*-sinapic acid grown by evaporation from a mother liquor containing no protein (see section 2.3). Compare these crystals with figure 3 to determine their crystal axis orientation. The full width of the micrograph shown corresponds to a real distance of 1.2 mm.

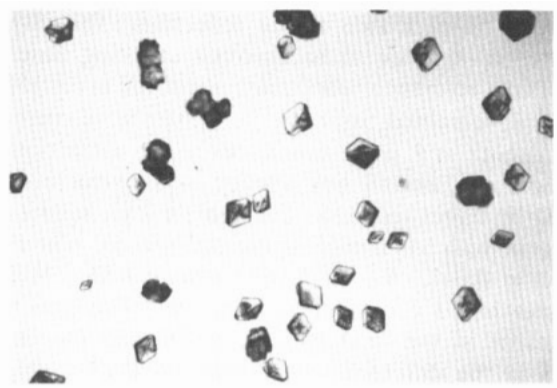


**Figure 5.** Crystals of *trans*-sinapic acid grown from a mother liquor with a high protein concentration (see section 2.3). Crystals with a hexagonal profile are oriented such that the $(10\bar{3})$ plane is in the plane of the photomicrograph. The full width of the micrograph shown corresponds to a real distance of 1.2 mm.

of other known protein matrix materials would help elucidate the mechanism of protein incorporation.

## 2. Experimental procedure

### 2.1. X-ray crystallography

*Trans*-sinapic acid crystals were grown by slow evaporation. A mother liquor was prepared using 2:1 water/-acetonitrile (v/v), 0.1% trifluoroacetic acid as the solvent. Five millilitres of the solvent were placed in a 1 x 10 cm glass test tube with an excess of *trans*-sinapic acid (Aldrich Chemical Company, catalogue number D13,460-0, lot number KY 07201TV). The mixture was placed in a bath held at 40 °C and stirred for 10 hours. The mother liquor was then decanted into another test tube, covered and replaced in the bath (50 °C) for 1 h without stirring. The covering was then replaced by a cap with a one millimetre diameter hole and the bath temperature decreased to 30 °C. The crystals that

formed at the surface of the mother liquor were removed by decanting the solvent. The crystal chosen for x-ray analysis was a colourless, monoclinic prism, with the dimensions $0.36 \times 0.21 \times 0.15$ mm. The crystal growth was performed in the dark, because *trans*-sinapic acid is known to be photoreactive in solution.

Data collection was performed on a Rigaku AFC6S diffractometer with graphite monochromatized MoK$\alpha$ radiation ($\lambda = 0.071069$ nm). Cell constants and an orientation matrix for data collection were obtained by least squares using the setting angles of 23 reflections ($29.1° < 28° < 38.°$). Data collection used the $\omega$–$2\theta$ scan technique. Omega scans of several intense reflections made before collection had an average width at half height of 0.43° with a take-off angle of 6.0°. Scans of $(1.26 + 0.30 \tan\theta)°$ were made at a speed of 8° min$^{-1}$ (in $\omega$). The weak reflections ($I < 10.0\sigma(I)$) were rescanned (maximum of two rescans) and the counts were accumulated to ensure good counting statistics. Stationary background counts were recorded on each side of the reflection. The ratio of peak counting time to background counting time was 2:1. Three reference reflections, measured every 150 reflections, remained constant. Intensities were corrected for Lorentz and polarization effects; a correction for secondary extinction was applied. The structure was solved by direct methods. Full matrix least squares refinement with all non-H atoms anisotropic converged to $R = 0.047$, $R_W = 0.033$, GOF= 1.65. Weighting was based on counting statistics. The maximum shift/error in the final cycle was 0.00 and the largest peaks in the difference map were $(+)0.18 \times 10^3$ and $(-)0.19 \times 10^3$ e nm$^{-3}$. Atomic scattering factors were from the *International Tables for X-ray Crystallography*; anomalous dispersion effects were included. All calculations were made with the TEXSAN software package.

### 2.2. Protein inclusion during crystal growth

The crystals shown in figure 3 were grown from a mother liquor composed of 2:1 water/acetonitrile (v/v), 0.1% trifluoroacetic acid, 6 $\mu$M horse skeletal muscle myoglobin (Sigma Chemical Company, catalogue M-0630, lot 61H7100) and 100 $\mu$M Coomassie Brilliant Blue G-250 that was saturated with *trans*-sinapic acid at 20 °C. The mixture was centrifuged and the supernant decanted into a $1 \times 10$ cm glass test tube. The tube was placed in a refrigerator (4 °C) and the crystals formed overnight.

The effect of sodium dodecylsulfate on protein inclusion was tested by growing crystals from the same mother liquor as above, but with 0.1% (w/w) of sodium dodecylsulfate added.

### 2.3. Effect of protein inclusion on crystal habit

The crystals shown in figure 4 were grown from a mother liquor containing 2:1 water/acetonitrile (v/v) and 0.1% trifluoroacetic acid, saturated with *trans*-sinapic acid at 20 °C. Approximately 5 $\mu$l of this solution were allowed to dry in room air on a glass microscope slide. The crystals shown in figure 5 were produced in the same manner, but the mother liquor contained 60 $\mu$M horse skeletal muscle myoglobin.

## 3. Results

### 3.1. X-ray crystallography

The crystal structure of sinapic acid has been determined and the x-ray diffraction data submitted to the Cambridge Crystal Structure Database. Some important structural parameters are given in table 1. (The atoms referred to in table 1 are labelled in figure 1.) The crystals were monoclinic, with $P2_1/n$ symmetry with $Z = 4$ and grew as hexagonal prisms, with the long axis in the [100] direction. The lattice parameters were $a = 0.4760(1)$ nm, $b = 1.5686(4)$ nm, $c = 1.4185(3)$ nm and $\beta = 90.30(2)°$ and the crystal had a calculated density of 1.406 g cm$^{-3}$.

Figure 2 shows the $(10\bar{3})$ plane of the crystal, which forms the end pinacoid of the prisms. *Trans*-sinapic acid forms a two-dimensional hydrogen-bonded polymer in which the doubly hydrogen-bonded carboxylic acid dimers are linked by hydrogen bonding between the *para*-hydroxyl group and the carboxyl group of adjacent molecules (see inset in figure 2). Adjacent sheets were not hydrogen bonded together. Only hydrogen participating in hydrogen bonds were shown. The hydrogen-bonding protons were located from difference maps.

### 3.2. Protein inclusion during crystal growth

Crystals of *trans*-sinapic acid were produced in a solution containing a protein-specific dye to determine which of the crystal faces were responsible for protein inclusion during growth. The dye used was Coomassie Brilliant Blue G-250, a non-covalent anionic dye that attaches to positively charged proteins. A non-covalent dye was chosen so that the protein would not be significantly modified by the presence of the dye agent. If no protein was present in the mother liquor, the dye was not incorporated into crystals: i.e. the crystals were clear. When a small amount of protein was present, dye was incorporated in hourglass zones, shown in figure 3. These zones start at a point inside the crystal (presumably the seed site) and extend through the crystal as flattened cones that terminate on the external $(10\bar{3})$ face. Etching these crystals showed that the zones were not external artifacts. The stained zones were dichroic, with the polarization axis perpendicular to the $(10\bar{3})$ end pinacoid. This property allowed the observation of the zones with low protein concentrations. It was necessary to grow large crystals to observe the zoning because the tablet crystal habit was thin in the [010] direction relative to its other dimensions. High protein concentrations in the mother liquor could not be used to intensify the staining in smaller crystals, because of the inclusion inhibition process illustrated in figures 4 and 5.

**Table 1.** Positional parameters for *trans*-sinapic acid in lattice units, and isotropic temperature factors ($\times 10^{-2}$ nm$^2$).

| Atom | x | y | z | $B_{eq}$ |
|---|---|---|---|---|
| O(1) | 0.2549(7) | 0.0193(2) | 0.0784(2) | 4.1(2) |
| O(2) | -0.0433(7) | 0.1136(2) | 0.0126(2) | 4.1(2) |
| O(3) | 0.4175(7) | 0.5523(2) | 0.1548(2) | 4.9(2) |
| O(4) | 0.7964(6) | 0.5191(2) | 0.2859(2) | 4.3(2) |
| O(5) | 0.9249(6) | 0.3617(2) | 0.3397(2) | 3.9(2) |
| C(1) | 0.165(1) | 0.0936(3) | 0.0682(3) | 3.3(3) |
| C(2) | 0.304(1) | 0.1633(3) | 0.1198(3) | 3.6(3) |
| C(3) | 0.240(1) | 0.2454(3) | 0.1102(3) | 3.1(3) |
| C(4) | 0.389(1) | 0.3157(3) | 0.1575(3) | 2.8(2) |
| C(5) | 0.324(1) | 0.3992(3) | 0.1321(3) | 3.2(3) |
| C(6) | 0.461(1) | 0.4676(3) | 0.1751(3) | 3.3(3) |
| C(7) | 0.663(1) | 0.4521(3) | 0.2446(3) | 3.3(3) |
| C(8) | 0.725(1) | 0.3673(3) | 0.2696(3) | 2.9(3) |
| C(9) | 0.592(1) | 0.3001(3) | 0.2270(3) | 3.1(2) |
| C(10) | 0.198(1) | 0.5722(3) | 0.0894(4) | 5.2(3) |
| C(11) | 0.995(1) | 0.2775(3) | 0.3716(3) | 4.2(3) |
| H(1) | 0.1872 | -0.0246 | 0.0443 | 4.4 |
| H(1A) | -0.1038 | 0.0761 | -0.0068 | 4.6 |
| H(2) | 0.4500 | 0.1488 | 0.1626 | 4.2 |
| H(3) | 0.0878 | 0.2598 | 0.0696 | 3.7 |
| H(4) | 0.9174 | 0.5050 | 0.3257 | 4.9 |
| H(5) | 0.1833 | 0.4098 | 0.0859 | 4.0 |
| H(9) | 0.6391 | 0.2431 | 0.2441 | 3.7 |
| H(10A) | 0.1862 | 0.6319 | 0.0814 | 6.3 |
| H(10B) | 0.2383 | 0.5460 | 0.0302 | 6.3 |
| H(10C) | 0.0246 | 0.5508 | 0.1123 | 6.3 |
| H(11A) | 1.0613 | 0.2444 | 0.3197 | 5.0 |
| H(11B) | 1.1358 | 0.2801 | 0.4187 | 5.0 |
| H(11C) | 0.8315 | 0.2506 | 0.3965 | 5.0 |

The effect of adding sodium dodecylsulfate to the mother liquor was to eliminate crystal staining by the dye. No change in the crystal habit was produced for the detergent concentrations used.

### 3.3. Effect of protein inclusion on crystal habit

Figure 4 shows crystals grown by drying a drop of saturated *trans*-sinapic acid solution. The dominant crystal habits were thin tablets and long prisms. The tablets in figure 4 have a large (010) face (in the plane of the photograph) and prominent $(10\bar{3})$ pinacoids. Figure 5 shows the crystals obtained when a high concentration of protein was present in the mother liquor. The crystal habit has changed to thin hexagonal prisms, with the $(10\bar{3})$ face now in the plane of the photograph.

## 4. Discussion

### 4.1. X-ray crystallography

The crystal structure obtained for *trans*-sinapic acid is analogous to those known for 4-hydroxycinnamic (*para*-coumaric [10]), 3,4-dihydroxycinnamic (caffeic [11]), and 2,5-dihydroxybenzoic (gentisic [12]) acids. These materials also form carboxyl-to-carboxyl dimers and arrange themselves into hydrogen-bonded sheets, forming crystals of $2/m$ point group symmetry and either $P2_1/a$ (gentisic acid), $P2_1/c$ (*para*-coumaric acid) or $P2_1/n$ (caffeic acid) space groups. The *trans*-sinapic acid crystals have the flattest sheet structure of the four. The structure in figure 2 shows one hydrogen attached to each oxygen in the carboxylic acid head group of the molecules. This feature was caused by the inability of the software used to assign an unambiguous bond for the hydrogens on the basis of difference mapping.

It is tempting to suggest that there is a correlation between the closely related space group symmetries of the best matrix materials and their ability to bind proteins. This type of speculation is probably unfounded, because the space group $P2_1/c$ is the most common space group for organic molecules [13]. Also sinapic acid is on the borderline between monoclinic and orthorhombic space groups, suggesting that some orthorhombic crystals may also have the appropriate properties for protein binding. The following section suggests that a specific geometrical feature of these structures is responsible for protein binding rather than their space group, i.e. the presence of flat, polarizable, hydrogen-bonded sheets.

### 4.2. Protein inclusion during crystal growth

The presence of the hourglass zones in figure 3 is characteristic of inhomogeneous inclusions [14]. The zoning pattern observed demonstrated that the protein was in-

cluded on the $(10\bar{3})$ face. The lack of incorporation on the other faces suggests that an extended region of the $(10\bar{3})$ plane was necessary for protein attachment: a kink in the growing face was not sufficient. The intensity of the staining was roughly proportional to the amount of protein present in the mother liquor. Bovine serum albumen and chicken egg white lysozyme were also tested, giving very similar results to those shown in figure 3.

The presence of low concentrations of sodium dodecylsulfate (known to inhibit protein ion production [14]) does not affect the *trans*-sinapic acid's crystal habit, but abolishes crystal staining. This observation implies that the detergent changes the interaction of the protein with the $(10\bar{3})$ face. The presence of sodium dodecylsulfate in solution with proteins is known to result in rod-like particles, composed of a protein molecule complexed with many detergent molecules [15]. The hydrophobic tail of the detergent molecules bind to the hydrophobic portions of the protein. The resulting particle no longer has the amphiphilic character of the original protein, suggesting that the amphiphilic nature of the protein is important for its interaction with the growing matrix crystal. (Note: amphiphilic molecules have geometrically separate hydrophilic and hydrophobic portions.)

The structure of the $(10\bar{3})$ face suggests the type of interactions that may be responsible for protein attachment during growth. It lacks free hydroxyl groups for hydrogen bonding. The methoxy group oxygens are free to act as hydrogen bond receptors, although space filling models show that the oxygens are sterically hindered from forming hydrogen bonds in most directions. The non-polar nature of most of the surface suggests that a 'hydrophobic' interaction may be responsible for protein attachment, i.e. the attachment of the protein to the surface results in an increased total entropy of the solvent/protein/crystal surface system. The protein attachment may be enhanced by the high $\pi$-electron density of this plane, resulting in a relatively polarizable surface. Permanent dipoles in the protein near the surface will be attracted to it by induced charge distributions.

The idea that hydrogen bonding is not involved in the protein inclusion process is strengthened by the nature of the crystal faces that do not bind proteins. They are all covered with moieties capable of forming hydrogen bonds (carboxylic acid and hydroxyl groups), but this does not appear to result in protein attachment. It is unlikely that hydrogen bonding could result in protein inclusion from an aqueous solution, because of the reduction in entropy produced by the formation of localized hydrogen bonds at a surface without any reduction in free energy. Evidence obtained from studies of protein crystallization and subunit assembly point to the importance of hydrophobic effects in the stabilization of protein–protein aggregate structures in aqueous solvents [16–18]. Here we propose that the same stabilization can be expected for the protein–surface–solvent interaction between an amphiphilic protein and the $(10\bar{3})$ face of sinapic acid crystals.

### 4.3. Effect of protein inclusion on crystal habit

The hypothesis that the $(10\bar{3})$ face was responsible for protein inclusion was supported by observations of crystal habit changes in solutions with different protein concentrations. The crystals shown in figure 4 resulted from the extension of the crystal in the [100] and [010] directions from the seed, leading to extended tablets with $(10\bar{3})$ pinacoids and an asymmetrically developed hexagonal prism. The presence of protein in the mother liquor produced a sharp decrease in the extension of the crystal in the [100] direction, allowing the faces of the prism to become more symmetrical (figure 5). This change in habit suggests that protein in the mother liquor has slowed the growth of the $(10\bar{3})$ face by blocking the growing pinacoid with a covering of bound protein molecules. This type of inclusion interference with the growth of a particular face is generally referred to as inclusion inhibition [19].

## 5. Conclusions

Using the information obtained by x-ray crystallography, the structural element of sinapic acid crystals responsible for including proteins is the extended hydrogen-bonded sheet of sinapic acid molecules in the $(10\bar{3})$ plane. It is proposed that proteins bind to this surface because of their amphiphilic nature, i.e. the hydrophobic portions of the protein become associated with the relatively hydrophobic $(10\bar{3})$ face. High concentrations of protein in the mother liquor have the effect of altering the sinapic acid crystal habit. Comparison of the crystal structure of sinapic acid with the known structures of gentisic, caffeic and *para*-coumaric acid shows that these materials have planes that correspond to the $(10\bar{3})$ plane of sinapic acid. This similarity leads to the prediction that it is these planar structures that result in protein binding in the other materials.

## Acknowledgments

RCB would like to thank Professor Roger Mason (Department of Earth Sciences, Memorial University) for valuable discussions about the nature and nomenclature of crystal inclusions. RCB would also like to thank Lorne Taylor (Department of Biochemistry, Memorial University) for providing some of the chemicals used in this study.

## References

[1] Karas M and Hillenkamp F 1988 *Anal. Chem.* **60** 2299–301
[2] Beavis R C and Chait B T 1989 *Rapid Commun. Mass Spectrom.* **3** 233
[3] Beavis R C and Chait B T 1991 *Chem. Phys. Lett.* **181** 479
[4] Chan T-W D, Colburn A W, Derrick P J, Gardiner D J and Bowden M 1992 *Org. Mass Spectrom.* **27** 188

[5] Strupat K, Karas M and Hillenkamp F 1991 *Int. J. Mass Spectrom. Ion Processes* **111** 89
[6] Beavis R C and Chait B T 1990 *Proc. Nat. Acad. Sci.* **87** 6873
[7] Beavis R C and Chait B T 1989 *Rapid Commun. Mass Spectrom.* **3** 432
[8] El-Bayouni S Z and Neish A C 1966 *Phytochem.* **5** 683
[9] Alibert G, Marigo G and Boudet A 1968 *CR Acad. Sci. Paris D* **267** 2144
[10] Bryan R F and Forcier P G 1980 *Mol. Cryst. Liq. Cryst.* **60** 157
[11] Garcia-Granda S, Beurskens G, Beurskens P T, Krishna T S R and Desiraju G R 1987 *Acta Crystallogr.* **C43** 683
[12] Haisa M, Kashino S, Hanada S-I, Tanaka K, Okazaki S and Shibagaki M 1982 *Acta Crystallogr.* **B38** 1480
[13] Haisa M 1978 *Acta Crystallogr.* **A34** 753
[14] Hillenkamp F, Karas M, Beavis R C and Chait B T 1991 *Anal. Chem.* **63** 1193A
[15] Reynold J A and Tanford C 1970 *J. Biol. Chem.* **245** 5161
[16] Frey M, Genovesio-Taverne J-C, Fontecilla-Camps J C 1991 *J. Phys. D: Appl. Phys.* **24** 105
[17] Littlechild J A 1991 *J. Phys. D: Appl. Phys.* **24** 111
[18] Janin J, Miller S, Chothia C 1988 *J. Mol. Biol.* **204** 155
[19] Buckley H E 1951 *Crystal Growth* (New York: Wiley)