Acta Cryst. (1998). D**54**, 598–609

Atomic Resolution Structure of Human HBP/CAP37/Azurocidin

Solveig Karlsen,[a]*† Lars Fogh Iversen,[b] Ingrid Kjøller Larsen,[a] Hans Jacob Flodgaard[b] and Jette Sandholm Kastrup[a]

[a]*Department of Medicinal Chemistry, Royal Danish School of Pharmacy, Universitetsparken 2, DK-2100 Copenhagen, Denmark, and* [b]*Novo Nordisk A/S, Novo Allé, DK-2880 Bagsvaerd, Denmark.*
E-mail: solveig.karlsen@chem.uit.no

(*Received 18 August 1997; accepted 10 November 1997*)

Abstract

Crystals of human heparin binding protein (HBP) diffract to 1.1 Å when flash-frozen at 120 K. The atomic resolution structure has been refined anisotropically using *SHELXL*96. The final model of HBP consists of 221 amino-acid residues of 225 possible, three glycosylation units, one chloride ion, 15 precipitant ethanol molecules and 323 water molecules. The structure is refined to a final crystallographic *R* factor of 15.9% and $R_{\rm free}$(5%) of 18.9% using all data. A putative protein kinase C activation site has been identified, involving residues 113–120. The structure is compared to the previously determined 2.3 Å resolution structure of HBP.

1. Introduction

The heparin binding protein (HBP),‡ also known as CAP37 or azurocidin, is an inactive serine protease homologue (Flodgaard *et al.*, 1991). The inactivity is caused by selective mutations (H41S and S175G) in the active-site triad. The human HBP exhibits high sequence identity (44%) with human neutrophil elastase (HNE) (Flodgaard *et al.*, 1991). The major 28 kDa form of HBP, isolated from human neutrophils, consists of 222 amino acids and is glycosylated *via* three putative *N*-glycosylation sites, which are all assumed to be occupied. The purified recombinant HBP (expressed in the baculovirus system) used in this study carries a three-residue C-terminal extension (Gly-Pro-Ala) and is a 27 kDa protein, corresponding to a glycosylation degree of 13% (Rasmussen *et al.*, 1996).

HBP and HNE might stem from a common ancestor, but during evolution HBP has adopted new functions in host defence against invading organisms by mechanisms not involving protease activity. HBP seems to have adapted a central role in both the first and second wave of defence in inflammation. It has been shown that HBP affects endothelial and fibroblast cells to contract (Ostergaard & Flodgaard, 1992), and that HBP stimulates the protein kinase C (PKC) activity in endothelial cells (Pereira *et al.*, 1996), suggesting that HBP is involved in the diapedesis of neutrophil leukocytes. Overlapping peptides derived from the primary structure of HBP were synthesized to investigate the structural basis for the PKC stimulation activity. The three most potent peptides found were peptide 95–122, peptide 140–165, and peptide 38–53 (listed in rank of potency). It was concluded that the PKC stimulation activity of HBP most likely could be related to the most potent peptide 95–122 (Pereira *et al.*, 1996). However, it remains to be shown whether the PKC stimulation activity is coupled to an endothelial cell contraction or detachment. The neutrophil leuko-

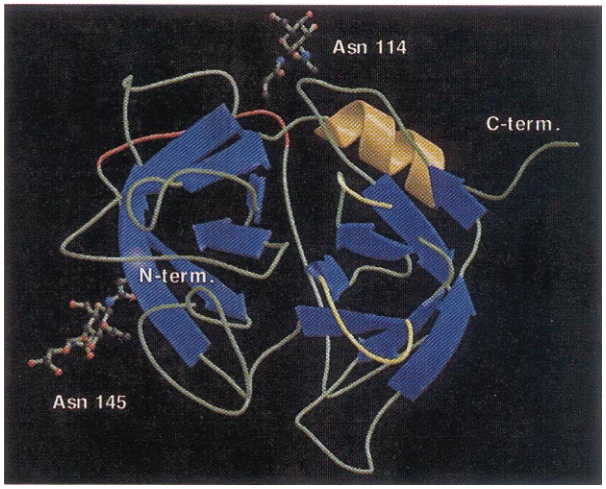


Fig. 1. A ribbon drawing of the 1.1 Å HBP structure. The β-strands are in blue, the α-helix in orange, and the loops in green. The putative lipid A binding site (residues 20–26 and 41–43) and the putative protein kinase C binding site (residues 113–120) are shown in yellow and red, respectively, and the glycosylation sites (Asn114 and Asn145) are marked. Four residues (44–47) have not been located (shown as a gap in the ribbon). The drawing was made by the use of *Rasmol* (Sayle, 1994), *MOLSCRIPT* (Kraulis, 1991), and *Raster3D* (Merritt & Murphy, 1994).

† Present address: Department of Chemistry, Faculty of Science, University of Tromsø, N-9037 Tromsø, Norway.

‡ Abbreviations: CAP37, cationic antimicrobial protein of 37 kDa; HBP, heparin binding protein; HNE, human neutrophil elastase; Fuc, fucose; GlcNAc, *N*-acetylglucosamine; Man, mannose; PKC, protein kinase C.

cyte migration to the infection site is also facilitated by HBP-mediated fibroblast contraction (Ostergaard & Flodgaard, 1992) and neutrophil leukocyte elastase-mediated digestion of extracellular matrix proteins. HBP and HNE are released from the same neutrophil leukocyte azurophile granules.

At the infection site HBP participates in killing the bacteria. HBP possesses potent antimicrobial action that

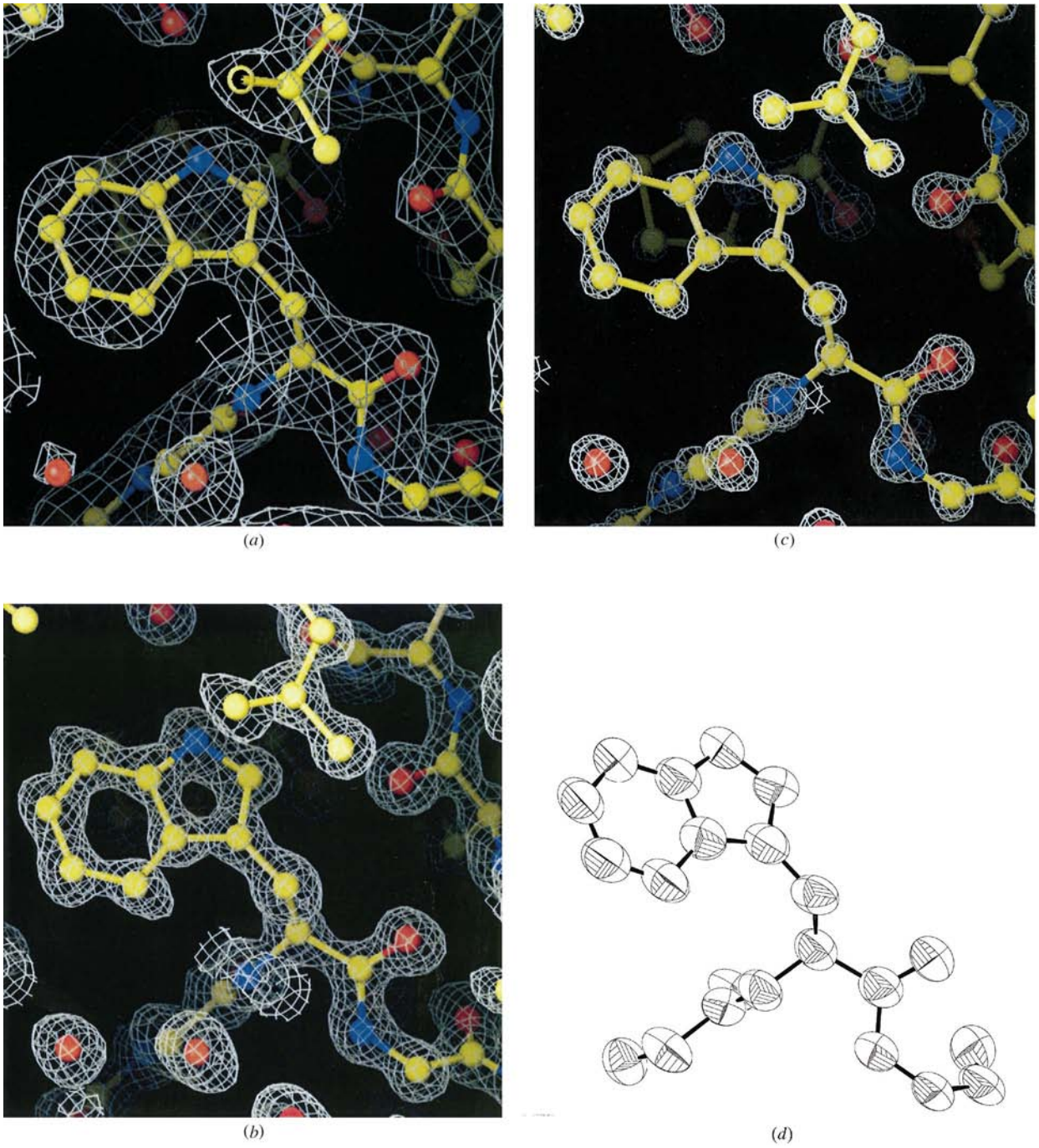


Fig. 2. Electron-density maps ($2F_o - F_c$) around Trp128. The figures (*a*)–(*c*) were produced using *O* (Jones & Kjeldgaard, 1993): (*a*) contoured at 1.5σ after isotropic refinement of the 2.3 Å HBP structure in *TNT*; (*b*) contoured at 1.5σ after isotropic refinement of the 1.1 Å HBP structure in *TNT*; (*c*) contoured at 8σ after anisotropic refinement of the 1.1 Å HBP structure in *SHELXL*; (*d*) An *ORTEP* representation (Johnson, 1976) of HBP in the same view as (*c*). The anisotropic displacement ellipsoids are drawn at the 50% probability level (for non-H atoms).

Table 1. *Survey of the data-collection statistics for the 1.1 Å data set*

$R_{\text{merge}}(I) = \sum_h \sum_i |I_{hi} - \langle I_h \rangle| / \sum_h \sum_i |\langle I_h \rangle|$, where I_{hi} is the intensity of the ith measurements of the same reflection and $\langle I_h \rangle$ is the mean observed intensity for that reflection.

Resolution range (Å)	Unique reflections	Completeness (%)	R_{merge} (I)	$\langle I/\sigma(I) \rangle$
20.00–3.04	4803	89.8	0.037	13.79
3.04–2.41	5034	98.0	0.047	19.24
2.41–2.11	4264	84.3	0.053	16.12
2.11–1.91	3703	73.2	0.047	14.02
1.91–1.78	3876	77.0	0.060	11.03
1.78–1.67	4726	94.7	0.070	10.40
1.67–1.59	4722	94.3	0.086	8.37
1.59–1.52	4699	94.5	0.103	6.98
1.52–1.46	4707	94.4	0.130	6.09
1.46–1.41	4684	94.1	0.151	5.25
1.41–1.37	4661	94.3	0.180	4.67
1.37–1.33	4646	93.7	0.210	4.27
1.33–1.29	4641	93.1	0.238	3.77
1.29–1.26	4589	93.0	0.275	3.36
1.26–1.23	4585	92.5	0.304	3.25
1.23–1.21	4556	91.7	0.311	3.01
1.21–1.18	4553	92.5	0.360	2.69
1.18–1.16	4543	92.5	0.387	4.45
1.16–1.14	4492	90.8	0.438	2.19
1.14–1.12	4055	81.9	0.479	1.96
20.00–1.12	90539	90.5	0.058	7.13

Table 2. *Progress in the SHELXL refinement of the heparin binding protein*

Round	Parameter	Restraint	R†	R_{free}‡	$R(4\sigma)$§	R_{free} (4σ)	Notes
1	8020	6936	23.5	25.4	21.03	23.40	Isotropic refinement of the whole structure, eight ethanols and 279 waters
2	8307	7299	23.2	24.9	20.7	22.7	Isotropic refinement, disorder modelled, deletion of waters
3	8291	7019	23.0	24.5	20.4	22.10	Isotropic refinement, deletion of more waters and five sugar rings
4	16269	19259	20.0	22.5	17.5	20.0	Anisotropic refinement of protein. Isotropic refinement for sugars, ethanols (12) and waters (240)
5	16929	19845	19.2	22.4	16.7	19.9	Anisotropic refinement of protein and sugars. Isotropic refinement for chloride ion, ethanols (17) and waters (275)
6	18472	21505	18.0	21.5	15.5	19.1	Anisotropic refinement of protein, sugars, ethanols (17) and 200 waters. Isotropic refinements of 10 waters
7	18741	21780	17.5	21.4	15.0	18.9	Anisotropic refinement of protein, sugars, ethanols and 263 waters. Isotropic refinement of 68 waters
8	18822	21905	16.7	20.5	14.2	18.0	Including riding H atoms. Anisotropic refinement of the whole structure, 16 ethanols and 323 waters
9	18777	21741	16.5	20.3	14.0	17.8	Anisotropic refinement of the whole structure, editing waters
10	18867	21930	15.9	18.9	13.4	16.4	Anisotropic blocked-matrix refinement of the whole structure, 15 ethanols ansd 323 waters

† R factor = $\sum \big||F(o)| - |F(c)|\big| / \sum |F(o)|$. ‡ *R factor calculated for a random 5% of the observations omitted from the refinement process.* § $R(4\sigma)$ R factor calculated for $>4\sigma(F_o)$.

is preferentially directed against Gram-negative bacteria (Shafer *et al.*, 1984, 1986; Campanelli *et al.*, 1990). The bacteria-killing mechanism is probably initiated by binding of HBP to the lipid A component of endotoxin which is located in the outer bacterial membrane. Recently, we identified a putative lipid A binding site (see Fig. 1) involving residues Asn20, Gln21, Gly22, Arg23, Phe25, Cys26, Ser41, Cys42 and Phe43 of HBP (Iversen *et al.*, 1997).

HBP is shown to have high affinity for heparin found in the extracellular matrix (Flodgaard *et al.*, 1991). The surface electrostatic potential of HBP reveals a large patch of basic residues concentrated on one side and a fairly acidic patch on the other side of the HBP molecule (Iversen *et al.*, 1997). It seems likely that the negatively charged heparin molecule binds to the basic patch by non-specific electrostatic interactions.

HBP has been shown to be a potent chemoattractant for monocytes (Flodgaard *et al.*, 1991; Pereira *et al.*, 1990). HBP, continuously secreted from migrating neutrophil leukocytes, may form a haptotactic gradient that attracts peripheral monocytes to the injured tissue in the second wave of inflammation (Ostergaard & Flodgaard, 1992). Furthermore, HBP stimulates mono-

cyte survival and induces morphological changes in monocytes towards a macrophage phenotype (Ostergaard & Flodgaard, 1992). Recently, it was demonstrated that HBP also possesses chemotactic action towards T cells (Chertov *et al.*, 1996).

Previously, we have reported the crystallization and molecular replacement solution of HBP based on a 3.1 Å room-temperature data set (Iversen *et al.*, 1996) and an HBP structure refined against a 2.3 Å room-temperature data set (Iversen *et al.*, 1997). The HBP structure seen in Fig. 1 consists of two similar domains of the same fold: a closed β-barrel of six antiparallel β-strands, characteristic of the trypsin-like serine protease superfamily. All the β-strands are found in the structurally conserved regions of the HBP sequence based on the alignment of human and porcine HBP and nine serine proteases (Flodgaard *et al.*, 1991).

Here we report the 1.1 Å cryogenic temperature (120 K) structure of HBP in which a putative protein kinase C activation site is identified.

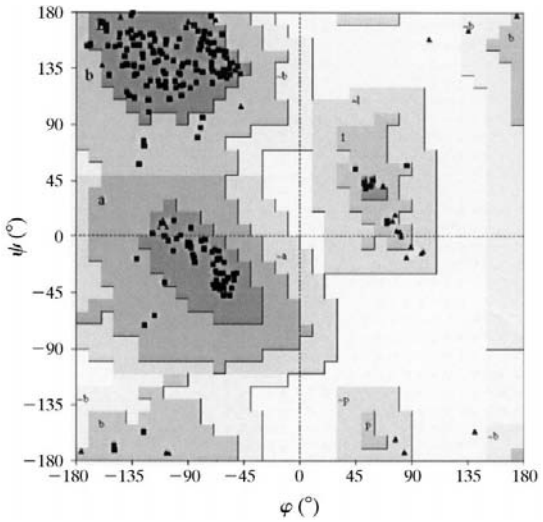


Fig. 3. Ramachandran plot of the final 1.1 Å HBP structure produced by using the program *PROCHECK* (Laskowski *et al.*, 1993). The triangles indicate the φ and ψ angles of the glycine residues and the squares the φ and ψ angles of the non-glycine residues. Levels, indicated by dark grey, medium grey and light grey shading represent regions which are most favoured, additionally allowed and generously allowed, respectively.

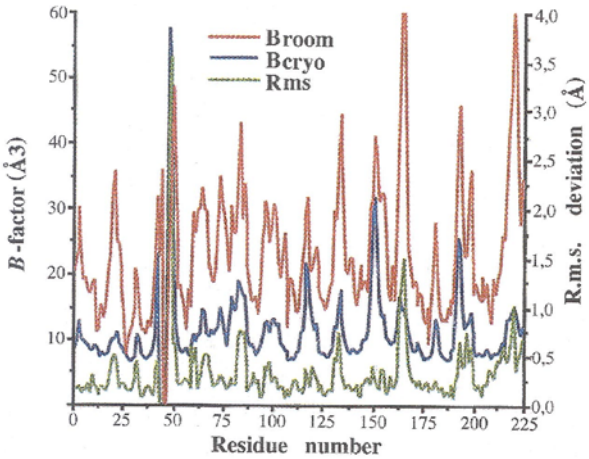


Fig. 4. Variation in average temperature factors of backbone atoms for the 2.3 Å room-temperature HBP structure (in red, B_{room}) and for the 1.1 Å cryo-temperature HBP structure (in blue, B_{cryo}). The r.m.s. deviations (R.m.s.) of backbone atoms for the two superimposed structures are shown in green.

2. Experimental

2.1. *Crystallization*

The HBP used for crystallization was expressed and purified as previously described (Rasmussen *et al.*, 1996). The crystals for data collection were grown with the repeated seeding technique using hanging drops containing 2 µl protein solution and 2 µl reservoir solution. The reservoir solution (1 ml) contained 14% ethanol, 5% 1,2,6-trihydroxyhexane, and 5 m*M* $MgSO_4$ in 0.1 *M* Tris–HCl buffer at pH 7.2. After one round of microseeding and macroseeding, respectively, the crystals were typically 0.4 × 0.4 × 0.3 mm in size. One crystal was flash-frozen for data collection at cryogenic temperature (120 K) using the loop method of Teng (1990). Approximately 20 s prior to freezing the crystal, 3 µl of 40% glycerol in water were added as cryoproctectant to the drop containing the crystal, giving rise to a total glycerol concentration of approximately 20% around the crystal.

2.2. *Data collection and processing*

Synchrotron data from the flash-frozen HBP crystal were collected on the X11 beamline in Hamburg to 1.1 Å resolution with a big MAR plate (300 mm) and a wavelength of 0.9165 Å. The crystal-to-detector distance was 130 mm, and each image was collected using a 0.4° oscillation range and an exposure time of approximately 140 s. A second round of data collection (low resolution) to 2.2 Å was performed on a small MAR plate with a crystal-to-detector distance at 200 mm, 0.8° oscillation, and exposure time of approximately 30 s. Both data sets were auto-indexed and processed with *DENZO* and scaled using *SCALEPACK* (Otwinowski & Minor, 1997). The 1.1 Å data set consists of 90 539 unique reflections which gives a completeness of 90.5% and an R_{merge} of 5.8% (see Table 1). The redundancy of the data is 2.7 (number of observations/number of unique reflections) and the crystal mosaicity is 0.29°. The cell dimensions for the orthorhombic unit cell ($P2_12_12_1$, one molecule per asymmetric unit) are a = 38.22, b = 65.16, and c = 101.58 Å which shows that the a and b axes

are slightly shorter for the cryogenic data set compared with the room-temperature data set (a = 39.19, b = 66.12 and c = 101.36 Å) (Iversen *et al.*, 1997). The overall temperature factor estimated from the Wilson plot (Wilson, 1942) is 8.8 Å^2 as compared with 37 Å^2 for the 2.3 Å room-temperature data.

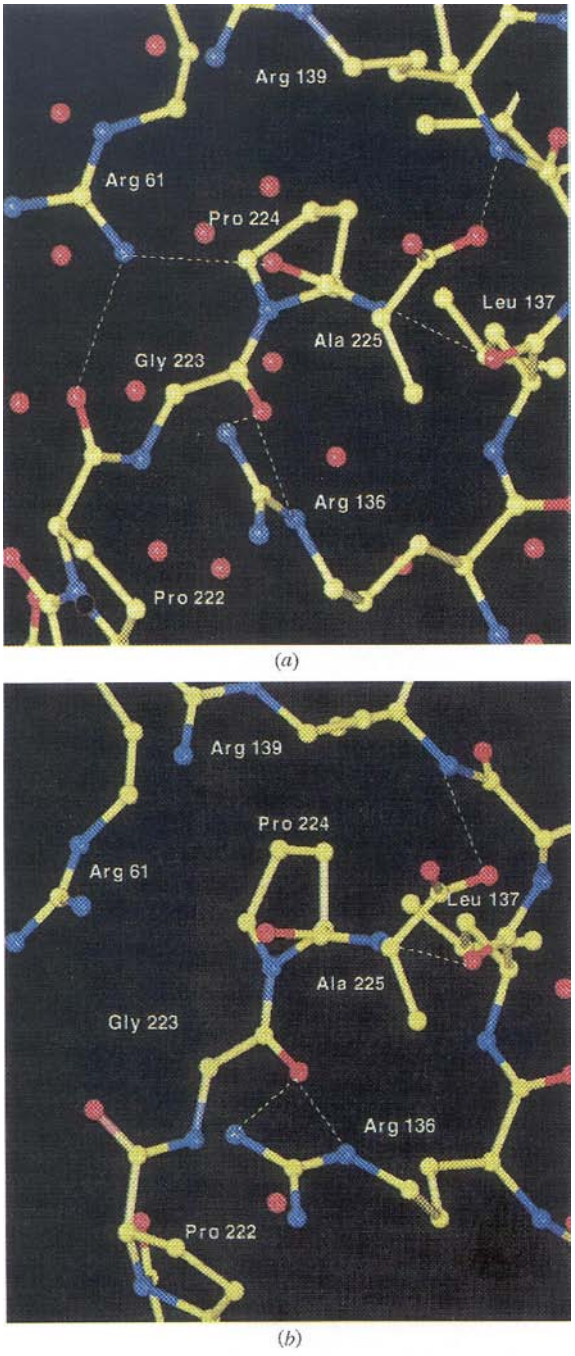


Fig. 5. Crystal packing of the C-terminal end of (*a*) the 2.3 Å HBP structure and (*b*) the 1.1 Å HBP structure. The intermolecular hydrogen bonds are marked as dotted lines. The fixation of Arg61 (by two hydrogen bonds) in the cryo-cooled HBP structure is clearly seen. The figure was produced by the program *O* (Jones & Kjeldgaard, 1993).

2.3. *Refinement*

The structure was refined in *SHELXL*96 (β-version) (Sheldrick & Schneider, 1997) after prior refinement in *TNT* (Tronrud, 1992) using the coordinates from a preliminary 2.8 Å HBP structure (Iversen, 1997). In both programs bond distances and angles were restrained using the standard values for Engh & Huber (1991). In *SHELXL* chiral volume restraints were applied to the Cα atoms (except for glycine) and to the Cβ atoms of threonine and isoleucine. Planarity restraints (e.s.d. 0.2 Å^3) were applied to the peptide units and aromatic side chains. In the anisotropic refinement, rigid-bond restraints were applied to the differences in mean-square displacement amplitudes along the 1,2- and 1,3-distances (e.s.d. 0.01 Å^2), and weak similarity restraints (e.s.d. 0.05 Å^2) were applied to the corresponding U_{ij} components of atoms which were close in space. The solvent water molecules were restrained weakly to be approximately isotropic. A diffuse solvent parameter was refined throughout, and anti-bumping restraints were generated automatically. The program *SHELXPRO* (Sheldrick & Schneider, 1997), which works as an interface between *SHELXL* and other macromolecular programs, was used for incorporating the coordinate file, crystal data, refinement instructions, and standard restraints into an input file for *SHELXL*96. The program was also used for converting the scaled data from *DENZO/SCALEPACK* (Otwinowski & Minor, 1997) to an *HKL*-file in *SHELXL* format with F^2 and σF^2 values and for calculating the electron density maps after the refinement rounds.

In both refinement programs atomic positions and temperature factors were refined by restrained least-squares minimization using the conjugate-gradient algorithm. Every refinement cycle consisted of least-squares minimization, calculation of $F_o - F_c$ and $2F_o - F_c$ maps, and manual adjustment of the model using the program *O* (Jones & Kjeldgaard, 1993). The data were divided into a working set (95% of data, 85 714 reflections) and a reference set (5% of data, 4511 reflections) with reflections randomly excluded from the full data set. The $R_{\rm free}$ data set was not identical for the *TNT* and *SHELXL* refinements. The reference set was used for monitoring the progress of refinement by calculation of $R_{\rm free}$ (Brünger, 1993). However, the progress of refinement may only be followed within each of the two different refinement programs as different $R_{\rm free}$ data sets were used. All data were used in the final refinement and map calculations, and in the final block refinement round.

After each refinement round, water molecules were added using information from the difference map and visually inspected. Ethanol molecules were included on the indication of the $F_o - F_c$ and $2F_o - F_c$ electron-density maps. One water molecule was replaced with a chloride ion because of high electron density (20σ level) and very low B value (3 Å^2). Modelling of disordered side chains and side chains with double conformations was carried out whenever the electron-density maps suggested these changes, and some of the sugar rings in the original HBP model were deleted due to lack of or diffuse electron density. Further refinement cycles were followed by addition of more water and ethanol molecules and deletion of some with high diffuse density and high temperature factors, fine tuning of disordered side chains, calculation of anisotropic thermal parameters and finally addition of riding H atoms. The H-atom positions were not refined, but calculated according to established geometrical criteria, and the temperature factors were set to 1.2 times the value of the atoms to which they were attached (1.5 for methyl and hydroxyl groups). Every new included water or ethanol molecule was first refined isotropically and finally anisotropically. The progress of refinement is shown in Table 2.

The program *PROCHECK* (Laskowski *et al.*, 1993) was used throughout the refinement and model building to investigate the stereochemistry of the structure and to point out regions needing further investigations.

The atomic coordinates and structure factors of the HBP structure have been deposited with the Protein Data Bank.†

Table 3. *Summary of statistics for the final set of HBP parameters*

No. of protein residues	221
No. of sugar rings	3
No. of water molecules	323
No. of ethanol molecules	15
No. of chloride ions	1
R.m.s. deviations from ideal values	
Bond distances (Å)	0.027 (0.03)
Angle distances (Å)	0.046 (0.06)
Planar groups (aromatic residues and main-chain peptides) (Å^3)	0.13 (0.20)
'Rigid bonds' $\lvert U_A^2 - U_B^2 \rvert$ (Å^2)	0.005 (0.010)
'Similar' U_{ij} values for spatially close atoms (Å^2)	0.038 (0.050)
Average temperature factors (Å^2)	
Protein (main chain)	11.4
Protein (side chain)	15.7
Saccharides	27.0
Ethanols	41.7
Waters	28.3
Chloride ion	10.2

3. Results and discussion

3.1. *Overall features*

The 1.1 Å data set has allowed building of a very precise model of HBP. The major differences between this HBP model and that previously reported (Iversen *et al.*, 1997) are largely a consequence of the low temperature and the higher limiting resolution (1.1 compared with 2.3 Å). The gain in resolution, going from room temperature to cryo-temperature, is obtained through a combination of reduction of thermal disorder, a longer life-time of the crystal due to reduced radiation damage, longer exposure time, and gentler crystal mounting. As diffraction was observed to 1.7 Å at room temperature for an HBP crystal before crystal decay (Iversen, 1997), diffraction was expected to 1.3 Å for a cryo-cooled crystal, based on the reduction in thermal disorder (Hope, 1988). The other factors mentioned above might contribute positively to enhance the diffraction to 1.1 Å.

The atomic resolution data have led to a much more extensive ethanol and water structure and clearer definitions of some previously poorly defined loops and side chains. Fig. 2 shows the improvement in electron density for the HBP structure by using 1.1 Å data (2*b*) and anisotropic refinement of B factors (2*c* and 2*d*). Addition of approximately 190 water molecules and several precipitant ethanol molecules has been possible compared to the 2.3 Å structure. Organic precipitant molecules from the crystallization assay do not usually give interpretable electron density. Dual side-chain locations of numbers of Ser and Met residues have been identified.

3.2. *Refinement of the HBP structure*

The structure of HBP was refined in *SHELXL*96 (β-version) (Sheldrick & Schneider, 1997) after prior refinement in *TNT* (Tronrud, 1992). The initial R factor was 38.2% for 95% of the data between 20.0 and 1.1 Å. Initially, five cycles of rigid-body refinements were performed lowering the R factor to 36.6%. The *TNT* refinement converged at an R factor of 22.3% of 221 protein residues, eight sugar rings, eight ethanol and 279 water molecules. The side chains modelled with two conformations were Met91 and Ser104.

Further isotropic refinement of the model was carried out in *SHELXL*. After three rounds of model building and refinement (see Table 2), the R factor converged at 20.4 and 23.0% for $F>4\sigma(F)$ and for all data, respectively. The model consisted now of 221 protein residues, three sugar rings, nine ethanol and 265

† Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1A7S, R1A7SSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: LI0274). At the request of the authors, the atomic coordinates and structure factors will remain privileged until 17 March 1999.

water molecules, and double conformations were built for Ser97 and Met76. As it was impossible to interpret electron density for five of the eight sugar rings, these were deleted from the structure. The R_{free} values dropped to 22.1 and 24.5% for $F>4\sigma(F)$ and for all data, respectively.

A significant drop in R factor was achieved when refining the HBP model anisotropically (see Table 2). In the second anisotropic run also the glycosylation sugar rings were refined anisotropically and two conformations were modelled for the side chain of Ser190. In this round of model building a chloride ion was substituted for a water molecule. In the subsequent refinements H atoms were included as riding atoms. Anisotropic refinement of 221 protein residues, three sugar rings, one chloride ion, 15 ethanol and 323 water molecules with data between 20.0 and 1.1 Å $[F>4\sigma(F)]$ converged with an R factor of 13.4% and an R_{free} value of 16.4%.

The final R factors are slightly higher than normally observed at this high resolution and for a protein of this size. One explanation for this might be found in the highly flexible glycosylation chains that cannot be fully interpreted in the electron-density maps. Only 17% of the glycosylation (determined by *MS*, Rasmussen *et al.*, 1996) has been included in the final model. Diffuse $F_o - F_c$ electron density (3–6σ level) is observed extending from the sugar rings, indicating that the more distal glycosylation atoms contribute to the structure factors, but cannot be built properly and refined in *SHELXL*. In addition, four residues (44–47) in a highly flexible loop cannot be seen in the electron-density maps. Another explanation might be a relatively low completeness of the data (84%) in the resolution range 20.0–8.0 Å as well as a reduced completeness in the resolution ranges 4.3–3.4 and 2.4–1.8 Å (Table 2) due to ice formation on the crystal. Furthermore, the R_{merge} in the high-resolution shells are increased as weak reflections are included. The R_{merge} drops to 17.2% in the resolution shell 1.14–1.12 Å using a 2.0σ cut-off. Therefore, the high final R factor is apparently not a result of the data quality at high resolution.

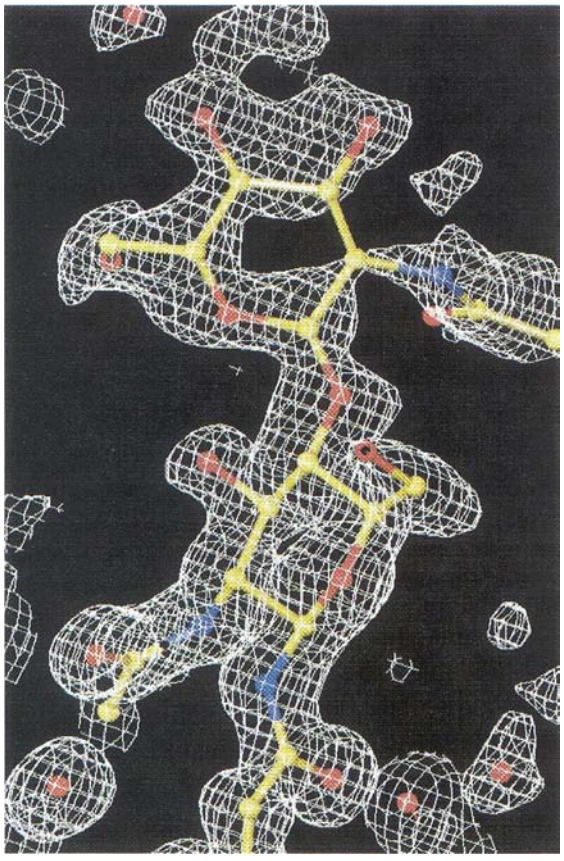


Fig. 6. Electron density $(2F_o - F_c)$ contoured at 3σ for the glycosylation moieties bound to Asn145. Two *N*-acetylglucosamine rings are built into the electron density. The figure was produced using *O* (Jones & Kjeldgaard, 1993).

3.3. *Quality of the HBP model*

The final model of HBP consists of 2091 atoms. The r.m.s. deviations from ideal values for bond lengths, planar groups and rigid bonds are given in Table 3. Luzzati plots (Luzzati, 1952) were produced using the program *SHELXPRO* (Sheldrick & Schneider, 1997) for the model after isotropic and anisotropic refinements. The mean error in atomic positions for the structures with isotropic and anisotropic temperature factors is 0.12–0.15 and 0.07–0.09 Å, respectively. The high accuracy of the final HBP model is confirmed by a Ramachandran plot as seen in Fig. 3 (Ramachandran & Sasisekharan, 1968) produced using the program *PROCHECK* (Laskowski *et al.*, 1993), showing that 90.3 and 9.7% of the residues are in the most favoured and in the allowed regions, respectively. Only 14 residues have $\chi1$ and $\chi2$ side-chain torsional angles deviating more than 2.5 standard deviations from ideal geometry. Ramachandran plots generated for the structures after isotropic refinement in *TNT* and *SHELXL*, show some additional residues in the allowed regions (11.9 and 10.8%, respectively) and fewer residues in the most favoured regions (88.1 and 88.6%, respectively) than the final structure with anisotropic temperature factors.

3.4. *Thermal parameters*

Analysis of the thermal parameters was carried out with the *SHELXPRO* program (Sheldrick & Schneider, 1997) after each cycle in *SHELXL*96. The average temperature factors for main-chain and side-chain atoms after isotropic refinement in *TNT* are 13.3 and 23.6 Å^2, and after anisotropic refinement in *SHELXL* 11.4 and 15.7 Å^2, respectively. This indicates that the temperature factors are overestimated in *TNT*

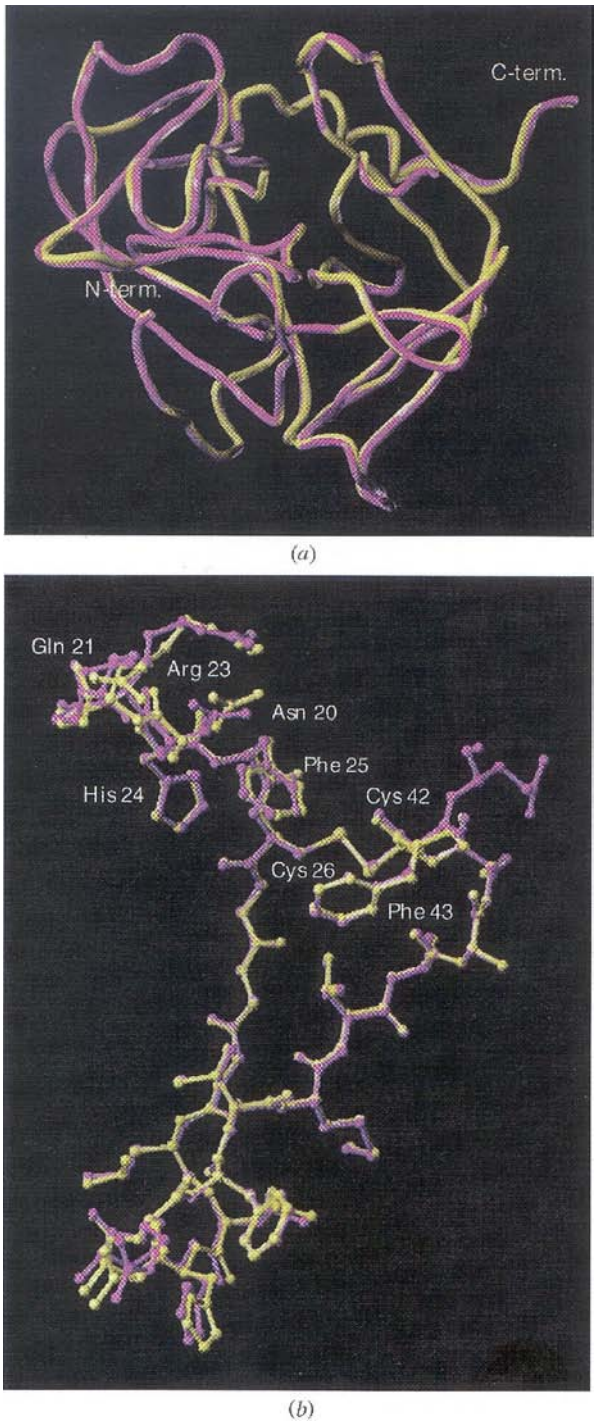


Fig. 7. (*a*) The superimposed backbone structures of HBP determined at 1.1 Å (yellow) and 2.3 Å (magenta) resolution. The tube representation was produced using *Sybyl* (Tripos Ass. Inc.). (*b*) A close up of the V-shaped sub-structure of HBP (residues 20–44). The putative lipid A binding site (residues 10–26 and 41–43) is located in the top of the V. The HBP structures refined to 1.1 and 2.3 Å are superimposed and coloured in yellow and magenta, respectively. The representation was produced using *Sybyl* (Tripos Ass. Inc.).

compared with *SHELXL*. The average temperature factors in the final HBP structure (see Table 3) for the chloride ion, saccharides, ethanol and water molecules indicate that the ethanols are more flexible than the waters.

A comparison (Fig. 4) of the temperature factors for the 2.3 Å HBP structure based on synchrotron data collected at room temperature and the high-resolution HBP structure based on cryogenic data, shows that the cryogenic structure is much less flexible. The main-chain and side-chain atoms of HBP to 2.3 Å resolution have average temperature factors of 23 and 35 Å^2, respectively, which is considerably higher than for the 1.1 Å structure (see Table 3). Thus, the $B_{\rm room}/B_{\rm cryo}$ ratio is 2.0 and 2.3 for main-chain and side-chain atoms, respectively.

Apart from the poorly defined 42–50 loop, all other loops seem to have reduced mobility in the 1.1 Å HBP structure. The flexibility is particularly reduced for the loops 20–26, 127–142, 163–178, and the C-terminal end of HBP, but one loop (149–157) has no significant reduction. In Fig. 4, the r.m.s. deviations in backbone atoms of the two superimposed structures are depicted, showing that the loop regions with high temperature factors are almost identical to the areas with structural differences.

As the C-terminal end in the 1.1 Å structure has clearly reduced mobility compared with the 2.3 Å structure, a detailed analysis of the intermolecular contacts in this area was performed (see Figs. 5*a* and 5*b*). The C-terminus of HBP is involved in important crystal packing contacts with a symmetry-related HBP molecule. Four of the hydrogen bonds are identical in the two HBP structures. The carbonyl O atom of Gly223 makes a bivalent hydrogen bond to NE and NH2 of Arg136. The other two interactions are seen from the carbonyl O and N atoms of Ala225 to the carbonyl O atom of Leu137 and the backbone N atom of Arg139, respectively. However, in the 1.1 Å HBP structure two extra hydrogen bonds are formed to the symmetry-related molecule. The carbonyl O atoms of Pro222 and Pro224 are hydrogen bonded to the NH1 atom of Arg61. It seems likely that these extra intermolecular contacts give the C-terminal end a tighter packing and hence reduced flexibility in the cryo-structure. Also, these results clearly show the efficiency of cryo-cooling in decreasing static and dynamic disorder in loop areas and terminal ends and pinpoint the advantage of this method when it can be used.

3.5. *The glycosylation sites*

Three *N*-glycosylation sites have been identified in the HBP sequence (Asn100, Asn114 and Asn145). An *N*-glycosylation site is recognised by the consensus sequence Asn-*Xaa*-Ser/Thr, where *Xaa* may be any amino acid (Lis & Sharon, 1993). All three glycosylation

sites of HBP have Thr in the consensus sequence (Thr102, Thr116 and Thr147). It has been reported that the most likely glycosylation of baculovirus expressed HBP is two $Man_3(Fuc)GlcNAc_2$ units and one $Man_3GlcNAc_2$ unit, assuming that all three sites are occupied (Rasmussen *et al.*, 1996). However, some microheterogeneity has been observed for HBP purified from baculovirus-mediated expression.

Well defined electron density for one *N*-acetylglucosamine ring at Asn114 can be observed in the final electron-density map at 1.1 Å resolution, and two *N*-acetylglucosamine rings are easily built into the electron density at the Asn145 (see Fig. 6). Additional density is observed for both glycosylation sites but has not been possible to interpret. There is no density for a proximal fucose at this glycosylation site, indicating that the glycosylation site lacks the fucose unit. This is consistent with the mass spectrum analysis performed by Rasmussen *et al.* (1996). Diffuse electron density for a proximal fucose can be seen for the glycosylation site at Asn114. If the glycosylation is $Man_3GlcNAc_2$ at Asn145 and $Man_3(Fuc)GlcNAc_2$ at Asn114, the glycosylation at Asn100 is $Man_3(Fuc)GlcNAc_2$.

For the third glycosylation site at Asn100, only limited density can be observed and no saccharide rings can be built with confidence. Also, the side chain of Asn100 has diffuse density. However, additional density in elongation with the ND2 atom indicates an *N*-glycosylation linkage. The side-chain atoms of Asn100 have relatively high temperature factors (26.7 $Å^2$), indicating that the glycosylation at this site is flexible. The glycosylation site is located at a large water channel spanning through the crystal and leaving plenty of space for movement of the flexible glycosylation chain.

3.6. *Solvent structure*

A total of 323 water molecules were included in the HBP structure which is 191 more than in the 2.3 Å structure. Regions of diffuse solvent were modelled using Babinet's principle which is included in the *SWAT* option in *SHELXL*96 (Moews & Kretsinger, 1975).

13 of the 15 ethanol molecules included in the final model of HBP are found at the surface of the protein and not in positions interferring with the protein folding. The other two ethanols, numbers 417 and 419, are partly buried in the HBP structure. Ethanol 417 forms hydrogen bonds to Gly56 and Ser67 and the ethanol 419 molecule is hydrogen bonded to Ala126 and Pro178. Three ethanol molecules (418, 422, and 431) are clustered on the surface of the molecule. Four molecules (420 and 430, 421 and 424) are located pairwise on the surface of HBP. The other six ethanol molecules are uniformly distributed over the HBP surface. Thus, the precipitant ethanol molecules do not interfere with the protein surface in any specific manner.

One chloride ion was located in the structure of HBP. The chloride ions were added to the crystallization assay through the 0.1 *M* Tris–HCl buffer. The chloride ion binds at the interface between two HBP molecules in a predominantly hydrophobic environment and it is likely that the chloride ion might play an important role in the crystal packing. The chloride ion is hydrogen bonded to two water molecules and to the backbone N atom of Leu112 at a distance of 3.4 Å. Other cases of chloride coordination to backbone N atoms have been observed. In dihydrofolate reductase a chloride ion is seen coordinating to two backbone N atoms with distances of 3.2 Å in an otherwise hydrophobic pocket (Brown *et al.*, 1993). Also, in rat mannose-binding protein a similar coordination of a chloride ion is found with distances of 3.0 and 3.1 Å (Kolatkar & Weis, 1996).

3.7. *Comparison of folding with the 2.3 Å HBP structure*

In Figs. 4 and 7(*a*), comparisons of the backbone-atom positions for the HBP structure refined against 2.3 and 1.1 Å data, respectively, are illustrated. The r.m.s. deviations for main-chain and side-chain atoms of the superimposed molecules are as low as 0.46 and 1.1 Å, respectively, indicating that the main chains of the two HBP structures are quite similar, but differences are found for some side chains. The figures show that the differences are exclusively found in the loop areas and in the C-terminus which also are determined to be the most flexible parts of the protein. The secondary-structure elements, α-helix and β-strands, are found at approximately the same locations in the two structures.

Particularly, the residues 48, 49, and 50 at the end of the highly flexible 44–48 loop have different conformations. In the 2.3 Å HBP structure the loop 44–48 was found to be flexible and these residues were not included in the final model. In the current 1.1 Å structure this loop still can not be properly traced. A sodium dodecyl sulfate–polyacrylamide gel electrophoresis gel run with an HBP crystal (identical to the crystal used for structure determination) shows only one band of 28 kDa (results not shown), indicating that no cleavage of the protein has occurred in this loop.

Differences in conformation are also found for the arginines at position 166–167 in a solvent-exposed loop and for the last two residues (224–225) in the structures. In addition, Arg61 has a different orientation in the cryo-cooled HBP structure due to the fixation of the C terminus (Fig. 5).

3.8. *Lipid A binding site*

It has been shown that HBP binds to the lipid A component of endotoxin which is the major component of the outer membrane of Gram-negative bacteria. A synthetic peptide corresponding to residues 20–44 of HBP mimics the antibacterial and endotoxin binding

action of the mature HBP (Pereira *et al.*, 1993). These residues fold into two antiparallel β-strands which span through the entire protein in a V-formed sub-structure as seen in Fig. 7(*b*). It was proposed earlier that the lipid A binding may involve the surface-exposed residues in the top of the V where Phe25, Cys26, Cys42 and Phe43 form a hydrophobic pocket suitable for binding of either fatty acid chains of glycosaminyl saccharides. A hydrophilic pocket is framed by Asn20, Gln21 and Arg23, suitable for binding of a glycosaminyl-linked phosphate group of lipid A (Iversen *et al.*, 1997).

A comparison of the HBP structures (Fig. 7*b*) refined to 2.3 and 1.1 Å, respectively, shows that few conformational changes are observed in the proposed binding area. The β-strands consist of the same residues, but small changes for the side chains are found in the connecting hairpin loop (31–34). In the putative lipid A binding site some small displacements are observed for the side chains, however, without affecting the overall features of the binding pocket.

The flexible loop (44–48) is positioned next to the putative lipid A binding site. It is proposed that the β(1–6)-linked GlcNAc backbone with two phosphoryl groups and not the aliphatic lipid chains binds to the protein (Iversen, 1997). The flexible loop may facilitate the binding of the flexible lipid A molecule. Remarkably, the corresponding loop in the HNE structure shows normal temperature factors for a loop region.

Difference electron density observed of an HBP:monophosphoryl lipid A crystal complex shows an *N*-acetylglucosamine ring bound as proposed within the hydrophobic pocket (Iversen, 1997). As no difference electron density, which could indicate binding to the lipid A binding site for the atomic resolution HBP structure, is observed, we conclude that the crystallized HBP contains no endotoxin that might have been bound during production and purification.

3.9. *Putative protein kinase C activation site*

An active or binding site of a protein is normally formed by a number of residues clustered in a certain area of the protein. The most potent PKC activation peptide (peptide 95–122) (Pereira *et al.*, 1996) runs on the surface of HBP, serving as a linker between the two domains of HBP and has no self-clustering residues (see Figs. 1 and 8*a*). It is not likely that all residues of peptide 95–122 take part in the PKC stimulation. PKC stimulating activity has not been reported for HNE (or other azurophile serine proteases), thus, the residues of HBP involved in the PKC activation might only include non-conserved residues compared with HNE. The non-conserved residues, based on the alignment previously reported, are 112–120 (see Fig. 8*b*) (Flodgaard *et al.*, 1991). The human and porcine HBP sequences are strictly conserved from residues 113 to 120 with a QNATVEAG sequence. The two HBP sequences have only Gln113 and Gly120 in common with the HNE sequence in this area. The strict residue conservation from residue 113 to residue 120 might imply a common PKC activation site across species of HBP. Also, the peptides used in the PKC activation study were generated from the human sequence and were found to be effective on rat endothelial cells (Pereira *et al.*, 1996). The synthetic peptides did not contain any glycosylation, indicating that the Asn114 glycosylation is not necessary for PKC activation. The residues 113–120 are located distant from both the basic patch and the putative lipid A binding site (Iversen, 1997). The site is lying on the edge of the acidic patch. Thus, HBP complexed with lipid A or heparin might stimulate PKC activity in endothelial cells.

Displaying the three most potent PKC activating peptides in the HBP structure (peptide 95–122, peptide 140–165, and peptide 38–53), a wide spread of the peptides is observed, see Fig. 8(*a*). The residues 113–120 and residues 146–159 show a clustering on one side of the HBP molecule (Fig. 8*b*). However, the human and porcine sequences do not contain any strict conservation of residues 146–159 as seen for residues 113–120. Thus, it is not convincing, that residues 146–159 are part of the PKC activation site, although there is clustering with residues 113–120.

The residues corresponding to the least potent of the three synthetic peptides (peptide 38–53) are overlapping with the residues of the putative lipid A binding site (Fig. 1) and are not in contact with the proposed PKC activation site, indicating that these residues are not part of the activation site. The observed PKC activity of peptide 38–53 and of peptide 140–165 might be explained by a Thr-Val sequence present in both peptides, mimicking the Thr116–Val117 sequence of the most potent peptide (peptide 95–122).

4. Concluding remarks

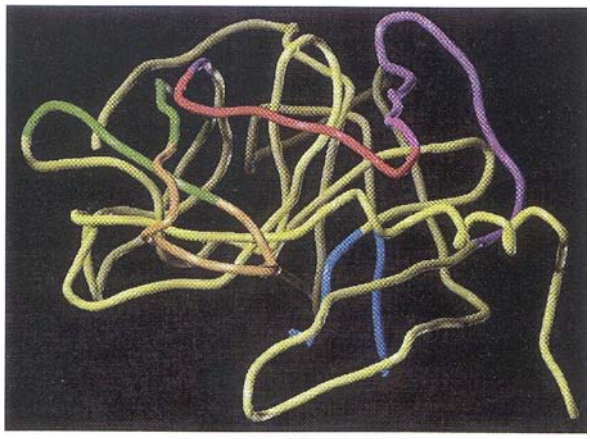
The atomic resolution data set to 1.1 Å of HBP has made it possible to refine the structure anisotropically and to build a very precise model of the protein. Several precipitant ethanol molecules and numerous water molecules and a chloride ion are included in the structure.

A comparison with the 2.3 Å room-temperature HBP structure shows that the structures are very similar. Minor differences are found in the loops which have become much less flexible in the cryo-cooled structure. This feature is seen particularly for the C-terminal extension of HBP, showing a more tight packing with a symmetry-related molecule in the 1.1 Å structure.

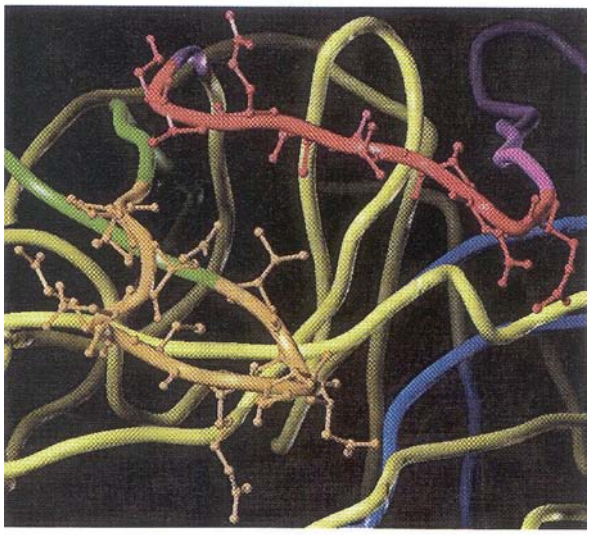
A putative PKC activation site of HBP includes residues 113–120 which are located on one side of the

protein apart from the putative heparin and lipid A binding sites. The residues 113–120 are strictly conserved between the human and porcine HBP. This implies that a common PKC activation site might exist among species.

This work was supported by grants from Pharma-Biotec, the Danish Research Academy, The Dansync Center for Synchrotron Radiation, and The EU Network Programme on Protein Crystallography, Contract Number CHRX-CT93–0143. We thank Brian Rosenberg for technical assistance and the European Union for support of the work at the EMBL Hamburg through the HCMP Access to Large Installations Project, Contract Number CHGE-CT93-0040.



(*a*)



(*b*)

Fig. 8. (*a*) A tube representation of HBP produced using *Sybyl* (Tripos Ass. Inc.). The three peptides showing PKC stimulating activity are 95–122, 140–165 and 38–53, listed in rank of potency. The conserved (95–112 and 121–122) and non-conserved residues (113–120) between HBP and HNE in peptide 95–122 are shown in magenta and red, respectively. In peptide 140–165 the conserved (140–148 and 160–165) and non-conserved residues (149–159) are shown in green and orange, respectively. The peptide 38–53 is shown in blue. The remaining structure is coloured in yellow. (*b*) A close up of the non-conserved residues, 113–120 and 149–159, in ball-and-stick representation. The colour coding is identical to that in part (*a*).

References

Brown, K. A., Howell, E. E. & Kraut, J. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 11753–11756.

Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.

Campanelli, D., Detmers, P. A., Nathan, C. F. & Gabay, J. E. (1990). *J. Clin. Invest.* **85**, 904–915.

Chertov, O., Michiel, D. F., Xu, L., Wang, J. M., Tani, K., Murphy, W. J. & Longo, D. L. (1996). *J. Biol. Chem.* **271**, 2935–2940.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Flodgaard, H. J., Ostergaard, E., Bayne, S., Svendsen, A., Thomsen, J., Engels, M. & Wollmer, A. (1991). *Eur. J. Biochem.* **197**, 535–547.

Hope, H. (1988). *Acta Cryst.* B**44**, 22–26.

Iversen, L. F. (1997). PhD thesis, Institute of Medicinal Chemistry, Royal Danish School of Pharmacy, Copenhagen, Denmark.

Iversen, L. F., Kastrup, J. S., Bjørn, S. E., Rasmussen, P. B., Wiberg, F. C., Flodgaard, H. J. & Larsen, I. K. (1997). *Nature Struct. Biol.* **4**, 265–268.

Iversen, L. F., Kastrup, J. S., Larsen, I. K., Bjørn, S. E., Rasmussen, P. B., Wiberg, F. C. & Flodgaard, H. J. (1996). *Acta Cryst.* D**52**, 1222–1223.

Johnson, C. K. (1976). *ORTEP*II. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.

Jones, T. A. & Kjeldgaard, M. (1993). *O The Manual, Version* 5.9.1, Uppsala, Sweden.

Kolatkar, A. R. & Weis, W. I. (1996). *J. Biol. Chem.* **271**, 6679–6685.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Laskowski, R., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lis, H. & Sharon, N (1993). *Eur. J. Biochem.* **218**, 1–27.

Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.

Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol.* **91**, 201–228.

Ostergaard, E. & Flodgaard, H. J. (1992). *J. Leukoc. Biol.* **51**, 316–323.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pereira, H. A., Erdam, I., Pohl, J. & Spitznagel, J. K. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 4733–4737.

Pereira, H. A., Moore, P. & Grammas, P. (1996). *J. Leukoc. Biol.* **60**, 415–422.

Pereira, H. A., Shafer, W. M., Pohl, J., Martin, L. E. & Spitznagel, J. K. (1990). *J. Clin. Invest.* **85**, 1468–1476.

Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Prot. Chem.* **23** 283–437

Rasmussen, P. B., Bjørn, S., Hastrup, S., Nielsen, P. F., Norris, K., Thim, L., Wiberg, F. C. & Flodgaard, H. J. (1996). *FEBS Lett.* **390**, 109–112.

Sayle, R. (1994). *Rasmol* 2.5: *A Molecular Graphics Tool*, Glaxo Research and Development, Greenford, Middlesex, UK.
Shafer, W. M., Martin, L. E. & Spitznagel, J. K. (1984). *Infect. Immun.* **45**, 29–35.
Shafer, W. M., Martin, L. E. & Spitznagel, J. K. (1986). *Infect. Immun.* **53**, 651–655.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Teng, T. Y. (1990). *J. Appl. Cryst.* **23**, 387–391.
Tronrud, D. E. (1992). *Acta Cryst.* A**48**, 912–916.
Wilson, A. J. C. (1942). *Nature (London)*, **150**, 151–152.