

第2章 放線菌 (*Streptomyces olivaceoviridis* E-86) 由来キシラナーゼの結晶構造解析

結晶化が成功しX線回折データが得られたら、構造解析の手順となる。類縁の蛋白質の立体構造が既に決定されており、その蛋白質との相同性が充分にあれば、その構造を利用して分子置換法により構造解析が可能となる。しかし、そうでない場合は、重原子同形置換法か多波長異常分散法による位相決定を行う必要がある。重原子誘導体結晶の調製に使われる代表的な重原子は、水銀、白金、金、ウラン、サマリウムなどであり、それぞれの原子について何種類もの試薬が試されている。重原子同形置換法では、結晶の同形性を保ったまま、重原子が結晶中にある程度の占有率をもって結合しなければならない。十分に位相決定するためには、重原子誘導体結晶データが複数あることが望まれる。本研究でも初めに重原子同形置換法を念頭に置き、重原子誘導体結晶の調製を試みたが、残念ながら有効な重原子誘導体を得られなかった。

1994年に構造解析されたGH10キシラナーゼの構造は、厳しい研究競争のため、Protein Data Bank (PDB) に登録された座標データはC α 原子のものであったり、公開がholdされていたりしており、分子置換法に適用できるものではなかった。1996年に相次いで *Pseudomonas fluorescens* (現在 *Cellvibrio japonicus*) (XYLA; Harris *et al.*, 1996), *Cellulomonas fimi* (Cex; White *et al.*, 1996) 由来キシラナーゼの触媒ドメインの結晶構造が更新され、座標データが使用できるようになった。FXYNの触媒ドメインとの相同性は30-40%であり、触媒ドメインについては分子置換法が応用できると考えられたため、構造解析に着手することにした。

2-1 方法

分子置換法

FXYN の調製、結晶化、データ測定は第 1 章に示したとおりである。分子置換法は、FXYN の全体構造に対する既構造決定モデルが存在しなかったため、触媒ドメインの構造のみをターゲットモデルとして分子置換法を試みることにした。まず、*Streptomyces lividans* XlnA (1xas) の触媒ドメインの C α モデル (Derewenda *et al.*, 1994) と *Cellulomonas fimi* Cex (2exo) の触媒ドメインの全構造 (White *et al.*, 1996) をもとに、プログラム QUANTA パッケージ (Accelrys) 中の MODELER を用い、FXYN の触媒ドメインの分子モデルを作製し、これを分子置換法のサーチモデルとした。分子置換法の回転関数、並進関数探索は CCP4 プログラムパッケージ (Collaborative Computational Project, 1994) 中のプログラム AMoRe を用いた (Navaza, 1994)。8 ~ 3.5 Å 分解能のデータを用いた回転関数探索では 2 つの突出した解が得られ、その相関係数はそれぞれ 0.125, 0.117 であった (Table 2-1)。まず、高い方の解に対し並進探索を行い、相関係数 0.150、R-値 53.4% の解を得た。本結晶は非対称単位中に 2 つの分子を含むので、最初に得られた解を固定し、2 番目に高い回転関数の解に対して更に並進探索を行った。その結果、相関係数 0.252、R-値 51.0% の解を得た。最終的に 10 ~ 3 Å 分解能のデータに対し解の精密化を行い、相関係数 0.521、R-値 46.6% という結果が得られた。

モデル構築と精密化

分子置換法の解を作製した触媒ドメインのモデルに適用し、2 分子の触媒ドメインの座標を作製後、プログラム X-PLOR (Brünger *et al.*, 1987) の simulated annealing を行い、そのモデルを

もとに電子密度図を作製したところ、モデルに一致した電子密度図が得られた。モデルの表示、修正はプログラム QUANTA (Accelrys) を用いた。 $2F_{\text{obs}}-F_{\text{calc}}$ 電子密度図、及び $F_{\text{obs}}-F_{\text{calc}}$ 電子密度図に従って、触媒ドメインモデルの修正を行い、プログラム X-PLOR による精密化を行うことを5サイクル繰り返し、その結果、2.5 Å 分解能での結晶学的 R-因子が 32.4%まで改善された。Free R-因子の計算には、データの約5%をランダムに抽出した (Brünger, 1992)。

十分に触媒ドメインのモデルを精密化した後、作製した F_o-F_c 電子密度図には XBD と思われる電子密度の塊が確認された。また、 $2F_{\text{obs}}-F_{\text{calc}}$ 、 $3F_o-2F_c$ 電子密度図で、触媒ドメインの近くに Trp 残基の電子密度図が確認された (Fig. 2-1)。この Trp 残基を中心にモデルを注意深く構築していき、モデルを加えながら位相の改良を進めていった結果、最終的に XBD の 313-436 番目のアミノ酸のモデルの構築に成功した。リンカー部分の電子密度図は最後まで観察されなかった。その後も、電子密度図を更新しながら、モデルの精密化、水分子の追加を行い、最終的に 1.9 Å 分解能のデータまで精密化を拡張した。モデルは、プログラム Procheck により確認し (Laskowski *et al.*, 1992)、グリシン以外のアミノ酸のうち 88.7%が Ramachandran 図 (Fig. 2-2) (Ramachandran & Sasisekharan, 1968) の most favorable 領域、11.3%が additional allowed 領域に帰属された。最終的な精密化の統計値を Table 2-2 に示す。

構造モデルの座標は PDB に登録した (PDB code 1xyf)。

2-2 結果

全体構造

FXYN の結晶構造を、分子置換法で 1.9 Å 分解能で決定した (Fig. 2-3) (Fujimoto *et al.*, 2000)。構造モデルは、結晶学的 *R*-因子 19.7%、Free-*R*-因子 24.2 %まで精密化された。最終モデルは、非対称単位中に 2 分子を含み、その 2 分子間の二乗根平均誤差は、触媒ドメインと XBD の両方の全原子に対して 0.54 Å と計算され、非対称単位中の 2 分子はほぼ同じ構造をとっているということが確認された。触媒ドメインは N 末端から 303 番目のアミノ酸まで、XBD については 313 番目から C 末端 436 番目までのアミノ酸のモデルを同定した。触媒ドメインは、他の GH10 キシラーゼと同様、 $(\beta/\alpha)_8$ -バレル (TIM-バレル) からなっており、これは多くの糖分解酵素によく見られる構造である。活性クレフトは、この TIM-バレルの中心にある β -バレルの C 末端側に存在し、触媒ドメインをほぼ 2 : 1 に分けるように全体に走っていた。XBD はサブドメイン α 、 β 、 γ の 3 つのサブドメインからなる 3 回繰り返し配列でできた球状ドメインであり、その構造は Murzin らによって提唱された β -trefoil 構造に類似していた (Murzin *et al.*, 1992)。両ドメインをつなぐリンカーは Pro/Gly リッチなものであり、そのうちの 9 残基については電子密度は見えず、構造は同定されなかった。

触媒ドメインの構造

触媒ドメインを作る TIM-バレルの β -ストランド、及び α -ヘリックスをそれぞれ C β 1-C β 8、C α 1-C α 8 と命名した (Fig. 2-3, 2-4)。Asp301 までが TIM-バレルの C α 8 に属するので、触媒ドメインは N 末端 1 残基目から 301 番目の Asp301 までとした。触媒ドメイン中には、TIM-バレ

ルの8つの α -ヘリックス以外に3つの α -ヘリックスが存在した。また、他のキシラナーゼと同様、*cis*-ペプチド結合がC β 3中 His81-Thr82の間に存在し、特殊な β -バルジ構造を形成していた。ジスルフィド結合は2つ存在し、Cys168-Cys201がC β 5-C β 6間を、Cys254-Cys260がC α 7-C β 8間をそれぞれ結ぶものであった。本酵素の触媒ドメインは、同じGH10キシラナーゼの中でも、*Penicillium simplicissimum*のキシラナーゼ (Schmidt *et al.*, 1998) とともに、301残基という最も短いタイプになる。

触媒活性クレフトは、TIM-バレルの中心にある β -バレルのC末端側に存在し、その中心にある2つのグルタミン酸、C β 4上のGlu128とC β 7上のGlu236が触媒残基であると考えられた (Fig. 2-5)。

キシラン結合ドメイン

XBDは、2つのストランドでできたヘアピンが6つ、あるいは α 、 β 、 γ の3つのサブドメインからなる3回繰り返し配列が、疑似3回軸周りに集まってできた球状ドメインであり、その構造はMurzinらによって提唱された β -trefoil構造に類似していた (Fig. 2-6, 2-7) (Murzin *et al.*, 1992)。この構造は疎水アミノ酸残基が中心に集まるようにできたシリンダー状の疎水コアをもっており、3つのサブドメインから集まった Trp 残基、Leu 残基、Ile (サブドメイン β では leu) 残基が上部層コアを、Leu、Leu (サブドメイン β では Ile)、Val (サブドメイン γ では Ile) が下部層コアを形成している (Fig. 2-8)。3つのサブドメインは同じ様な構造をとっており、サブドメイン間の二乗根平均誤差は、C α 原子に対して、1.75 ($\alpha - \beta$)、1.62 ($\beta - \gamma$)、および0.71 Å ($\alpha - \gamma$) と計算された。

この XBD の構造は、植物毒素のリシンやアブリン等のガラクトース結合レクチンとして知

られているもののB鎖とほぼ同じ構造である (Rutenber *et al.*, 1987; Tahirov *et al.*, 1995)。これらのレクチン蛋白質のB鎖は、3回繰り返し球状ドメインがさらに直列に2つ並んだ構造を持っている (Fig. 0-10)。XBD とこれらのレクチン構造を比較しても、疎水コアを中心に非常によく類似しており (Fig. 2-7)、アミノ酸配列の相同性は 20-30%程であるが、全体の二乗根平均誤差は、リシンB鎖の C α 原子に対して、1.27、1.68 Å、アプリンのB鎖の C α 原子に対して、1.15、1.40 Å と低く、全体の構造が類似していることが示された。

2-3 考察

触媒ドメイン

触媒ドメインのアミノ酸配列は他の GH10 キシラーゼと非常によく保存されている。糖分解酵素のサブサイトは通常、切断部位より還元末端の方に+1、+2・・・、非還元末端側に-1、-2・・・とサブサイト番号をつけるように統一されている (Davies *et al.*, 1997)。キシロオリゴ糖あるいはアナログとの複合体の構造が他の GH10 キシラーゼで構造決定されているので、そのうちのひとつである Cex/2-fruoro-2-deoxy-xylobiose (PDB code 2xyl, Notenboom *et al.*, 1998b) と重ねた図を作製したところ (Fig. 2-5)、サブサイト-2、-1、+1 の3つは非常によく保存されていることが分かった。即ち、サブサイト-1 は Asn127, His81, His207, Trp85, Trp266 及び Trp274 からなり、Trp266, Trp274, Lys48, Gln88, Glu44 及び Asn45 がサブサイト-2 を形成していると考えられる。このように一側のサブサイトはこれまで構造解析されているキシラーゼと比較しても、アミノ酸配列、及び、構造ともにほぼ完全に保存されていることが明らかとなった。また、サブサイト+1 と+2 はそれぞれ、Gln205, His207, Tyr172、および、Asn173, Asn209, Arg275 より構成されている。+側のサブサイトでは、サブサイト+1 はキシロースとスタッキングする

Tyr172 があり、この側鎖はほとんどのキシラナーゼで、Tyr 残基を持っているが、XYLA では Phe となっている。しかしながら、どちらにしても芳香環によるスタッキング結合にかかわることになり、よく保存されているといえることができる。それに対し、サブサイト+2 の周辺は Fig. 2-5 の左側において特に保存されていない。Charnock らはこの点について、Cex と XYLA を比較して、サブサイトの数の比較、キシランの認識機構の考察を行っている (Charnock *et al.*, 1998)。FYXN の場合も他のキシラナーゼと同様に、C β 6、C β 7 の C 末端側では特に違いが見られる。XYLA では2つの Tyr 残基が余分に存在し、サブサイト+3 や+4 の存在に関係しているように思われるが、FXYN のサブサイト+2 はほぼ溶媒領域に突出しており、その先にはサブサイトは存在しないと考えられた。

キシラン結合ドメイン

XBD の β -trefoil 構造を持つ3回繰り返し配列蛋白質は、そのアミノ酸配列の特徴から、(Gln-X-Trp)₃-ドメイン、あるいは Gly-X-X-X-Gln-X-Trp(Tyr)のキーコンセンサスモチーフを持つリシンスーパーファミリーとして以前から認識されている (Fig. 2-6) (Hazes, 1996; Hirabayashi *et al.*, 1998)。リシンの立体構造が明らかとなる以前から、Gln 残基が基質認識にかかわり、Trp 残基は疎水コア形成に必要であろうと考えられてきた。これらのファミリー CBM13 に属する蛋白質は、非常に類似した構造を持つにもかかわらず、基質特異性が異なるものが存在することが明らかとなっている。リシンの解析により、ガラクトース結合レクチンとして認識されてはきたが、シアル酸認識のレクチンや β -1,3-グルカンを認識するものなどが存在することも明らかになっており、我々の XBD は同様の構造を利用してキシランを認識するものである (Shimoi *et al.*, 1992; Kaku *et al.*, 1996)。

リシンのB鎖とラクトースの複合体の構造が決定されている (Fig. 2-6, 9)。リシンのB鎖は2つの β -trefoil 構造が直列に並んでおり、ドメイン1、2とすると、ドメイン1のサブドメイン α (サブドメイン1 α) とドメイン2のサブドメイン γ (サブドメイン2 γ) にラクトースが結合した構造が解明された (Rutenber *et al.*, 1987)。リシンB鎖の Asp22 または Asp234 及び、Asn46 または Asn255 がラクトースのガラクトースに水素結合しており、Gln47 または Gln256 は先程の Asn 及び Asp に水素結合していた (Fig. 2-6, 9)。また、Trp37 または Tyr248 の芳香環が、ちょうどガラクトースのヘキソース環にスタッキング結合していた。サブドメイン1 α では、Gln35 もラクトース結合に関与していた。これを XBD に当てはめてみると、これまであげたリシンのラクトース結合に関与しているアミノ酸は、XBD では厳密に保存されていた (Fig. 2-6)。リシンの構造と、XBD の構造を重ね合わせてみると、これらのアミノ酸は XBD でもリシンと同じ構造をとっていることが分かる。XBD のサブドメイン α 、 β 、 γ をリシンのラクトースと結合しているサブドメイン1 α と重ねてみると、全てのサブドメインでラクトースに結合しているアミノ酸が保存されているのが分かる (Fig. 2-9)。この結果、立体構造から判断する限り FXYN でもラクトースあるいはガラクトースに結合する可能性が高いと考えられた。

実際に、ガラクトースへの XBD の結合を久野らが測定した結果、XBD のサブドメイン α では、キシロースと同程度のガラクトース結合能が認められ、また、サブドメイン β 、 γ でもわずかながらガラクトースに結合することが明らかとなった (Kuno *et al.* 2000)。おそらく、XBD がガラクトースと結合する際はリシンがガラクトースと結合する様式と同じであると考えられるが、詳細な考察をするためには、FXYN とキシロオリゴ糖、あるいはラクトースとの複合体の立体構造解析が必要である。

XBD のキシラン結合部位は、芳香族アミノ酸の存在や、親水性アミノ酸が数多く存在する

ということなどから、リシンのガラクトース結合部位に対応するところであると予想された。しかし、リシンのラクトース結合構造をもとに XBD のキシラン結合構造モデルを作製すると、ガラクトースとキシロースの構造の相違によって立体障害が生じる。即ち、ガラクトースは 4 位の水酸基が axial に出ているのに対し、キシランでは equatorial に出ている点が異なるためである (Fig. 2-10)。また、キシランでは β -1,4-結合によりキシロースが直列に延びる構造をしているため、XBD がキシランに結合する際は XBD に結合したキシロースの 1 位及び 4 位の水酸基は結合部位より外に向くことのできる場所に来ると予想される (Fig. 2-9, 2-10)。従って、XBD のキシロース結合様式とリシンのガラクトース結合様式は異なると考えられる。

XBD とリンカーの役割

触媒ドメインと XBD の間のインタフェースは、主に親水性である。5 つの水素結合が、Ca7 の前半にある Ser243 から Ser245 と XBD の間にあり、他に水分子を介した水素結合も多数存在する (Table 2-3, Fig. 2-11)。触媒ドメインのこの領域は、FXYN の C β 7 と Ca7 が Cex と比較して短いので、XBD との間に立体障害のないユニークなインタフェースを形成する。この比較的しっかりとした結合様式の結果、マルチドメインを持つ FXYN の結晶化が成功したと考えられる。

XBD を取り除いて酵素活性を測定すると、XBD がある場合の方が無い場合よりも、不溶性キシランに対する加水分解能が高いことが明らかとなっているが、可溶性キシランに対しては XBD は影響を与えないことも示されている (Kuno *et al.*, 2000)。セルラーゼも XBD のようにセルロース結合ドメイン (CBD) を持っているものが多い。セルロースは通常、直鎖状である。

Thermomonospora fusca 由来のセルラーゼ E4 の構造は、セルロースの触媒部位と CBD のセルロ

ース結合部位が1直線上に並び、その間には、セルロースに結合あるいは支える芳香族アミノ酸が多く見られることが明らかとなっている (Sakon *et al.*, 1997)。即ち、直鎖状のセルロースが、セルロース結合部位から触媒部位まで直線的に伸びたまま結合し、それにより触媒反応は、CBDが触媒部位に基質を送り込むことで促進されることが示されている。

対照的に、XBDの糖結合部位とFXYNの触媒作用の触媒部位は、直線で結ぶことはできない。セルロースと比較してキシランは自由度が高く、安定化するときの構造は緩い3回らせん構造をとるといわれているが、不溶性キシランの場合は複雑な構造を持ち、ほとんど動くことはないと考えられる (Nieduszynski & Marchessault, 1972)。そこで、不溶性のキシランにうまく結合するために重要だと思われるのが、XBDがリンカーを通して触媒部位に対して溶液中で自由な動きをとることで、その動きによりキシランが両ドメインに同時に結合できると予想される。結晶中ではインターフェースを形成していた両ドメインは、溶液中では独立して存在すると考えた方がよいであろう。

FXYNの結晶構造では、サブドメイン β の基質結合部位が他のサブドメインより触媒クレフトの入口に非常に近い。しかし、リンカー部分の可動性によりドメイン同士が常に一定の構造をとっていないならば、結晶構造でのサブドメイン β の距離は特に問題にならず、他のサブドメイン α 、 γ においても基質に結合でき、その基質が触媒ドメインで分解される。結局、柔軟なリンカー部分により、XBDに3つの基質結合部位があることで基質に結合する機会が単純に約3倍になると考えられる。不溶性キシランが基質であるとき、これは特に効果的だと思われる。

XBDは、リシンスーパーファミリーと呼ばれるガラクトース結合レクチンの構造をとってはいるが、3回繰り返し配列により3カ所の基質結合部位を持つことにより、不溶性キシラン

に対してより結合しやすい構造を持っていること、および、フレキシブルなリンカー部分により、結合した基質を触媒ドメインに切断できるようにしているということが予想され、FXYNは非常に効率的に進化した酵素のひとつであると考えられる。

Table 2-1 Solutions of the rotation search, the translation search and the fitting procedure in the molecular replacement method calculated by the program AMoRe (Navaza, 1994).

Rotation solutions sorted by correlation coefficients

Best Solution: 21.118 52.098 337.396 12.543

| TABLE | Alpha | Beta | Gamma | Tx | Ty | Tz | Corr | Rfac |
|------------|-------|--------|-------|--------|---------|---------|---------|----------|
| SOLUTIONRC | 1 | 21.12 | 52.10 | 337.40 | 0.00000 | 0.00000 | 0.00000 | 12.5 0.0 |
| SOLUTIONRC | 1 | 134.77 | 53.52 | 336.30 | 0.00000 | 0.00000 | 0.00000 | 11.7 0.0 |
| SOLUTIONRC | 1 | 146.40 | 51.56 | 346.17 | 0.00000 | 0.00000 | 0.00000 | 9.1 0.0 |
| SOLUTIONRC | 1 | 105.05 | 60.61 | 127.34 | 0.00000 | 0.00000 | 0.00000 | 9.0 0.0 |
| SOLUTIONRC | 1 | 27.60 | 50.41 | 357.97 | 0.00000 | 0.00000 | 0.00000 | 8.9 0.0 |
| SOLUTIONRC | 1 | 26.50 | 50.77 | 1.00 | 0.00000 | 0.00000 | 0.00000 | 8.9 0.0 |
| SOLUTIONRC | 1 | 50.69 | 59.66 | 288.52 | 0.00000 | 0.00000 | 0.00000 | 8.6 0.0 |
| SOLUTIONRC | 1 | 69.14 | 42.23 | 201.22 | 0.00000 | 0.00000 | 0.00000 | 8.4 0.0 |
| SOLUTIONRC | 1 | 134.51 | 75.28 | 163.98 | 0.00000 | 0.00000 | 0.00000 | 8.4 0.0 |
| SOLUTIONRC | 1 | 39.80 | 65.42 | 174.70 | 0.00000 | 0.00000 | 0.00000 | 8.3 0.0 |

1st translation solutions sorted by correlation coefficients

| SOLUTIONTF1 | TABLE | Alpha | Beta | Gamma | Tx | Ty | Tz | Corr | Rfac | PkHtcounter |
|-------------|-------|-------|-------|--------|--------|--------|--------|------|------|-------------|
| SOLUTIONTF1 | 1 | 21.12 | 52.10 | 337.40 | 0.3798 | 0.3196 | 0.3334 | 15.0 | 53.4 | |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3798 | 0.3196 | 0.3334 | 15.0 | 53.4 | 1 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3826 | 0.1753 | 0.3324 | 12.5 | 54.3 | 3 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3810 | 0.2550 | 0.3380 | 11.9 | 54.0 | 4 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3800 | 0.3529 | 0.3324 | 11.7 | 54.3 | 7 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3801 | 0.0721 | 0.3335 | 11.2 | 54.2 | 8 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3794 | 0.2049 | 0.3322 | 10.8 | 54.3 | 6 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3814 | 0.1265 | 0.3329 | 10.7 | 54.6 | 2 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3735 | 0.0494 | 0.3333 | 10.4 | 54.6 | 5 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.3805 | 0.0114 | 0.3327 | 10.3 | 54.5 | 10 |
| SOLUT_1 | 1 | 21.12 | 52.10 | 337.40 | 0.2677 | 0.3040 | 0.2587 | 9.1 | 55.0 | 9 |

Table 2-1 continued.

2nd translation solutions sorted by correlation coefficients

SOLUTIONTF1 1 21.12 52.10 337.40 0.3798 0.3196 0.3334 15.0 53.4

SOLUTIONTF2 1 134.77 53.52 336.30 0.9757 0.1122 0.2852 25.2 51.0

| | TABLE | Alpha | Beta | Gamma | Tx | Ty | Tz | Corr | Rfac | PkHtcounter |
|---------|-------|--------|-------|--------|--------|--------|--------|------|------|-------------|
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.9757 | 0.1122 | 0.2852 | 25.2 | 51.0 | 1 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.4756 | 0.1123 | 0.1765 | 16.1 | 53.3 | 2 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.7060 | 0.0733 | 0.0377 | 15.9 | 53.7 | 4 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.8687 | 0.9448 | 0.1866 | 15.6 | 53.6 | 9 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.7704 | 0.9059 | 0.3683 | 14.8 | 53.8 | 3 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.9759 | 0.5590 | 0.7744 | 14.4 | 54.0 | 5 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.0302 | 0.5654 | 0.5514 | 14.3 | 54.0 | 8 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.4716 | 0.1129 | 0.3272 | 14.3 | 53.9 | 10 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.7162 | 0.9112 | 0.3670 | 14.1 | 54.1 | 7 |
| SOLUT_2 | 1 | 134.77 | 53.52 | 336.30 | 0.9985 | 0.1067 | 0.9026 | 14.0 | 53.7 | 6 |

Fitting Result

| | TABLE | Alpha | Beta | Gamma | Tx | Ty | Tz | Corr | Rfac |
|-----------|-------|--------|-------|--------|---------|---------|---------|------|------|
| SOLUTIONF | 1 | 21.66 | 52.06 | 336.23 | 0.38000 | 0.31795 | 0.33297 | 52.1 | 46.6 |
| SOLUTIONF | 1 | 133.85 | 53.38 | 339.80 | 0.98064 | 0.11405 | 0.28498 | 52.1 | 46.6 |

Table 2-2 Crystal parameter and refinement statistics.

| | | |
|--|---------|--------------------------|
| Cell parameter ($P2_12_12_1$) | | |
| a (Å) | | 79.8 |
| b (Å) | | 95.0 |
| c (Å) | | 141.1 |
| Resolution range (Å) | 8.0-1.9 | (1.97-1.90) ^a |
| No. of reflections in refinement | 75, 408 | (9,865) |
| Completeness (%) | 90.0 | (75.1) |
| R -factor ^b (%) | 19.7 | (30.5) |
| R_{free} -factor ^c (%) | 24.2 | (35.7) |
| No. of protein non-hydrogen atoms | | 6,470 |
| No. of water molecules | | 493 |
| Rmsd | | |
| Bond lengths (Å) | | 0.009 |
| Bond angles (°) | | 1.5 |
| Dihedral angles (°) | | 24.2 |

^a Values in parentheses are for the outermost shell.

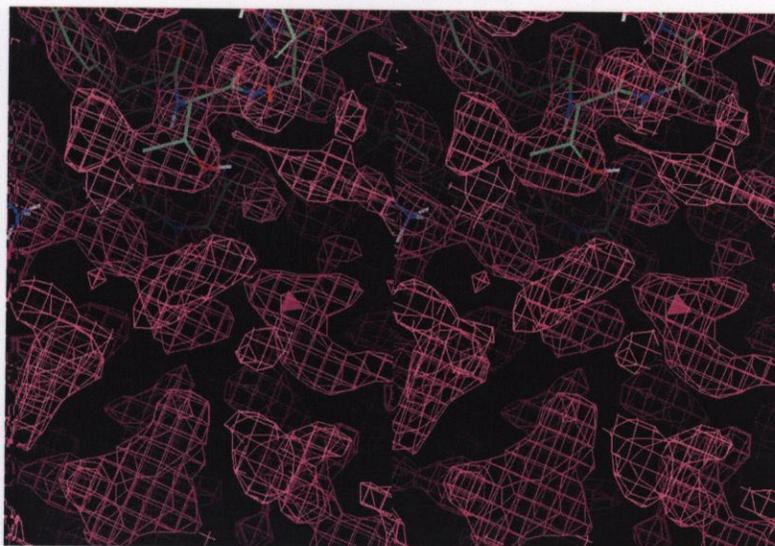
^b R -factor is defined as $R = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$.

^c R_{free} -factor was calculated using 5% of the unique reflections.

Table 2-3 Hydrogen-bonding and one-water-mediated interactions between the catalytic domain and XBD.

| Catalytic domain | Distance (Å) | | | XBD |
|------------------|--------------|------|-----|------------|
| OG Ser243 | 3.0 | | | OD2 Asp354 |
| N Ser244 | 3.1 | | | O Asp354 |
| OG Ser244 | 3.2 | | | O Asp354 |
| N Ser245 | 3.2 | | | OD1 Asp354 |
| OG Ser245 | 2.8 | | | OD1 Asp354 |
| OG Ser244 | 2.9 | W135 | 2.7 | O Thr353 |
| O Lys289 | 2.7 | W332 | 2.6 | NE Arg391 |
| O Lys289 | 2.7 | W332 | 2.6 | NH2 Arg391 |
| O Ala291 | 2.9 | W124 | 2.7 | OH Tyr352 |
| OG1 Thr294 | 2.6 | W252 | 3.3 | NH1 Arg391 |
| OG1 Thr294 | 2.6 | W252 | 2.6 | OG1 Asn393 |
| OD1 Asn298 | 3.0 | W141 | 3.1 | N Ser394 |

(a)



(b)

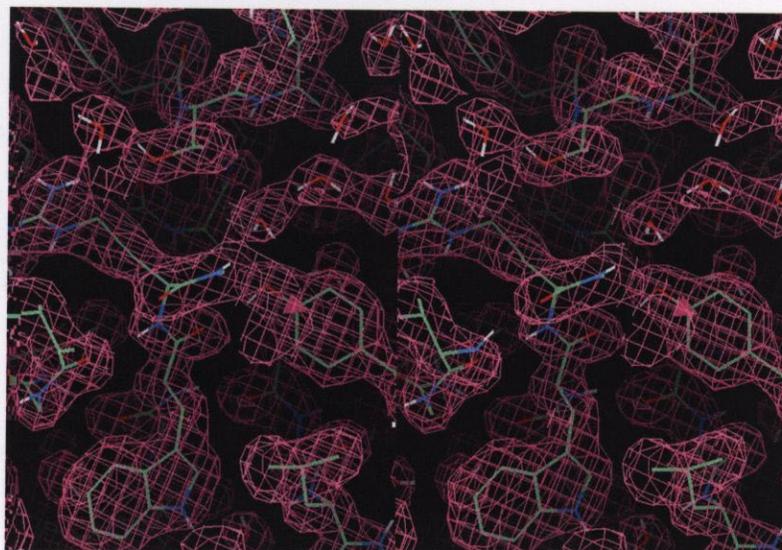
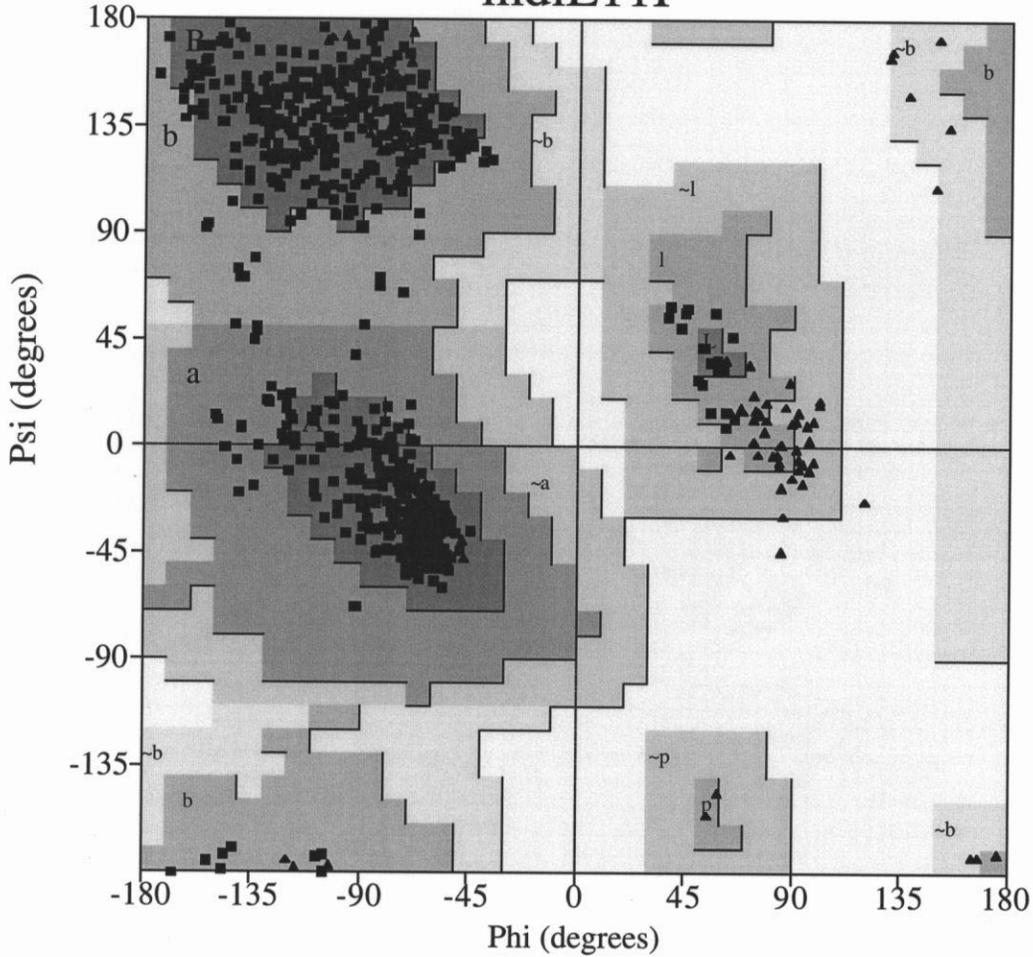


Fig. 2-1 Stereoview of the $2F_{\text{obs}} - F_{\text{calc}}$ electron density map around the interface between the catalytic domain and XBD contoured at 1σ . (a) Calculated before the tracing of the XBD, and (b) after the refinement of FXYN.

Ramachandran Plot

mdIE11F



Plot statistics

| | | |
|--|-----|--------|
| Residues in most favoured regions [A,B,L] | 651 | 88.7% |
| Residues in additional allowed regions [a,b,l,p] | 83 | 11.3% |
| Residues in generously allowed regions [~a,~b,~l,~p] | 0 | 0.0% |
| Residues in disallowed regions | 0 | 0.0% |
| ----- | | |
| Number of non-glycine and non-proline residues | 734 | 100.0% |
| Number of end-residues (excl. Gly and Pro) | 4 | |
| Number of glycine residues (shown as triangles) | 102 | |
| Number of proline residues | 14 | |
| ----- | | |
| Total number of residues | 854 | |

Based on an analysis of 118 structures of resolution of at least 2.0 Angstrom and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Fig. 2-2 Ramachandran plot produced by the program procheck (Laskowski *et al.*, 1992).

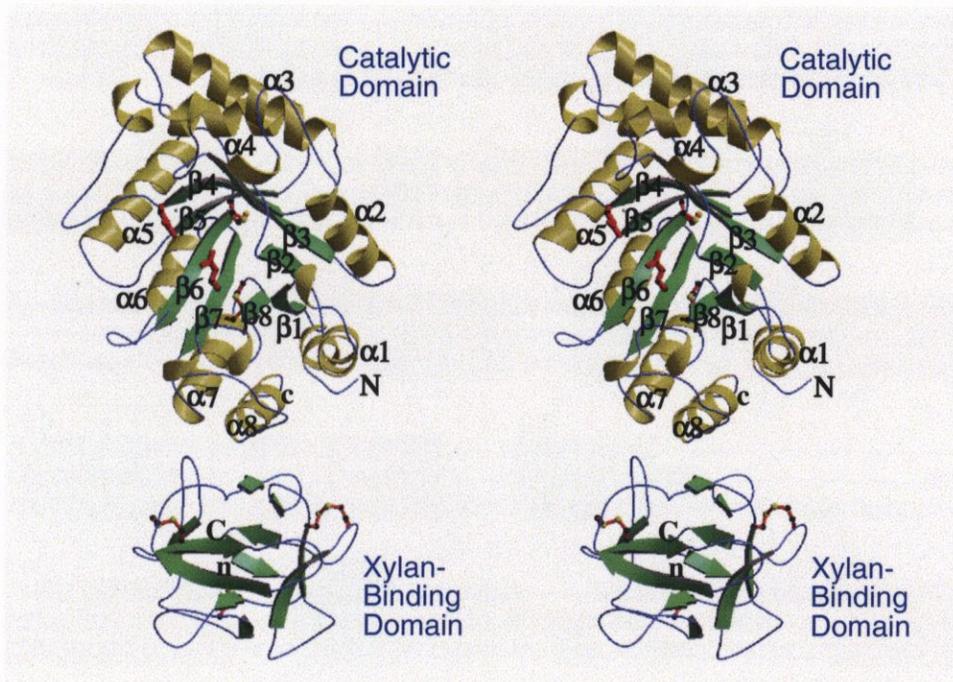


Fig. 2-3 Stereoview of the ribbon model of FXYN; catalytic domain (upper) and XBD (lower). β -Strands C β_1 -C β_8 and α -helices C α_1 -C α_8 build up the $(\beta/\alpha)_8$ -barrel of a catalytic domain and the two catalytic residues are displayed in red. The linker region lying between "c" and "n" is not shown. Disulfide bonds are indicated by ball-and-stick drawings. The figure was drawn with the program Raster3d (Kraulis, 1991; Merritt & Murphy, 1994).

| | | | | | | | |
|------|-----|---------------------------------|---|-------------------------------------|--|------------|-----|
| | | Cα0 | Cβ1 | Cα1 | Cβ2 | Cα2 | |
| FXYN | 1 | AESTLGA | AAQSGRYFGTAIASG | ----- | KLGDSAYTTIASREFNMVTAENEMKIDATEPQRGQFNFSAGDRV | | 67 |
| Cex | 1 | A-TTLKEAADG | AGRDFGFALDPN | ----- | RLSEAQYKAIADSEFNLVVAENAMKWDATEPSQNSFSFGAGDRV | | 66 |
| XYLA | 1 | ---GLASL--- | ADFP | IGVAVAASGGNADIFTSSARQNI | VRAEFNQITAENIMKMSYMYSG-SNFSFTNSDRL | | 65 |
| | | | | | | | |
| | | Cβ3 | Cα3 | Cβ4 | * | | |
| FXYN | 68 | YNWAVQNGKQVRGHTLAWHS | --QPGWMQSLSGSTLRQAMIDHINGVMGHYK | GKIAQWDVVNEAFSD | ----- | | 132 |
| Cex | 67 | ASYAADTGKELYGHTLVWHS | --QLPDWAKNLNGSAFESAMVNHVTKVADHFEGKVASWDVVNEAFAD | ----- | | | 131 |
| XYLA | 66 | VSWAAQNGQTVHGHALVWHPSYQLPNWASD | -SNANFRQDFARHIDTVAAHFAGQVKS | WDVVNEALFDSADDP | | | 136 |
| | | | | | | | |
| | | Cα4 | Cβ5 | Cα5 | | | |
| FXYN | 133 | ---DGSGRR-DSNLQRT | --GNDWIEVAFRTARAA | DPAAKLCYNDYNIENWTWAKTQGVY | NMVRDFKQ | RGVP | 198 |
| Cex | 132 | -----GGGRRQDSAFQOKL | -GNGYIETA | FRAARAADPTAKLCINDYNVEG | -INAKSNSLYDLVKDFKARGVP | | 196 |
| XYLA | 137 | DGRGSANGYR-QSVFYRQFGGPEYI | IDEAFRRARAADPTAELYNDFNTEE | -NGAKTTALVNLVQRL | LNNNGVP | | 206 |
| | | | | | | | |
| | | Cβ6 | Cα6 | Cβ7 | * | | |
| FXYN | 199 | IDCVGFQSHFNSG-SPYNSNFRTTLQNF | AALG--VDVAITELDIQ | ----- | ----- | G | 241 |
| Cex | 197 | LDCVGFQSHLIVG-QV-PGDFRQNLQRFADL | G--VDVRITELDIRMRT | ----- | ----- | PSDATKLA | 248 |
| XYLA | 207 | IDGVGFQMHVMNDYPS- | IANIRQAMQKIVALSP | TLKIKITELDVRLNNPYDGNSSNDY | TNRNDCAVSCAGLD | | 277 |
| | | | | | | | |
| | | Cα7 | Cβ8 | Cα8 | | | |
| FXYN | 242 | ASSSTYAAVTNDCLAV | ---SRCLGITVWGV | RDTDSW----- | RSGDT-PLLFNGDGSKKAAYTAVLNALN | | 301 |
| Cex | 249 | TQAADYKVVQACMQV | ---TRCQGVTVWGITDKYSWVPDVF | PGEA-ALVWDASYAKKPAYAAVMEAF | | | 312 |
| XYLA | 278 | RQKARYKEIVQAYLEV | VPPGRRGGITVWGIADP | SWL-YTHQNLPDWPLLFNDNLQPKPAYQGVVEALS | | | 345 |

Fig. 2-4 Topological sequence alignment of the catalytic domains of FXYN, Cex (White *et al.*, 1996) and XYLA (Harris *et al.*, 1996). The alignment was prepared based on the superposed models. Residues that belong to α -helices, α_{10} -helices and β -strands are indicated in orange, yellow and green, respectively. The secondary structure elements of the $(\beta/\alpha)_8$ -barrel are named as in Fig. 2-3. The catalytic residues are indicated with red asterisks.

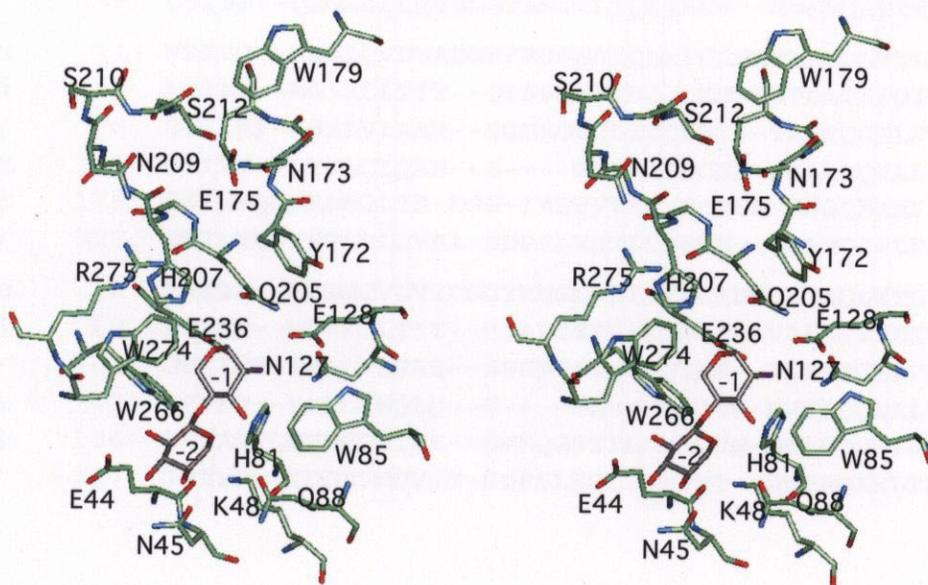


Fig. 2-5 Stereoview of the active cleft of FXYN with the inhibitor in the structure of Cex/2-fluoro-2-deoxy-xylobiose complex (2xy; Notenboom *et al.*, 1998) superimposed. Subsites -1 and -2 are labeled in the cleft.

| | | | | | | | | | | | | | | |
|-----------|-----|--------|----|-----------|----------|----------|--------|--------|---------|--------|--------|-------|-----|-----|
| | | | △ | ○△* | | △**△* | ○ | ** | △ | | | | | |
| XBDα | 313 | GQIKGV | - | GSGRCLDVP | NA | STTDGTQV | QLYDCH | -- | SATNQQW | TYTDA | 355 | | | |
| β | 356 | GELRV | -- | YGDKCLDAA | -- | GTGNGTKV | QIYSCW | -- | GGDNQK | WRLNSD | 395 | | | |
| γ | 396 | GSIVGV | - | QSGLC | LD | AVGGGTAN | GLIQLY | SCS | -- | NGSNQR | WTRT | 436 | | |
| Ricin B1α | 11 | VRIVG | -- | RNGLCVD | VRDGR | FHNGNAI | QLWPCK | SNTDAN | QLWTL | KRD | 54 | | | |
| 1β | 55 | NTIRS | -- | NG-KCL | TTY-- | GYS | PGVYV | MIYDC | N | AATD | ATRWQI | WDN | 95 | |
| 1γ | 96 | GTIINP | - | RSSLV | LAAT-- | SGNSG | TTLTV | QTNI | -- | YAVSQ | GWLP | TNN | 136 | |
| 2α | 142 | TTIVG | -- | LYGLC | LQAN-- | S---- | GQVW | IEDC | SS- | EKAE | QQWAL | YAD | 178 | |
| 2β | 179 | GSIRP | Q | QNRDN | CLTS- | DSN- | IRET | VVKIL | SCGPA- | SSG | QRW | FKND | 221 | |
| 2γ | 222 | GTILNL | - | YSGLV | LDVRA- | SDPSL | KQIIL | YPLH | -- | GDPN | QIWL | PLF | 262 | |
| Abrin B1α | 16 | VRIGG | -- | RDGMC | VDVYD | NGYHNG | NR | IIMW | KCKDR | LEEN | QLWTL | TKSD | 59 | |
| 1β | 60 | KTIRS | -- | NG-KCL | TTY-- | GYAPG | SYVMI | YDCT | S | AVAE | ATYWEI | WDN | 100 | |
| 1γ | 101 | GTIINP | - | KSALV | LSAE-- | SSSMG | GTLTV | QTNE | -- | YLMR | QGWRT | GNN | 141 | |
| 2α | 147 | TSISG | -- | YSDLC | MQAQ-- | G---- | SNVW | MAD | CDS- | NKKE | QQWAL | YTD | 183 | |
| 2β | 184 | GSIRSV | Q | NTN | NCLTSK-- | DHKQ | GSTIL | LMG | CSN- | GWAS | QRW | VFKND | 226 | |
| 2γ | 227 | GSIIYS | L- | YDDM | VMDV | KG- | SDPSL | KQIIL | WPYT | -- | GKPN | QIWL | TLF | 267 |

Fig. 2-6 Topological sequence alignment of XBD, with ricin B-chain (Rutenber & Robertus, 1991) and abrin B-chain (Tahirov *et al.*, 1995). The manner of partitioning the subdomains differs slightly from that in previous reports (Rutenber & Robertus, 1991; Tahirov *et al.*, 1995), in which the conserved Ile residues were positioned at the end. Residues in XBD that belong to α_{10} -helices and β -strands are indicated in yellow and green, respectively. Conserved Cys residues are marked by circles, residues in the hydrophobic core are marked by triangles, and residues corresponding to the sugar-binding site in the ricin B-chain are marked by asterisks.

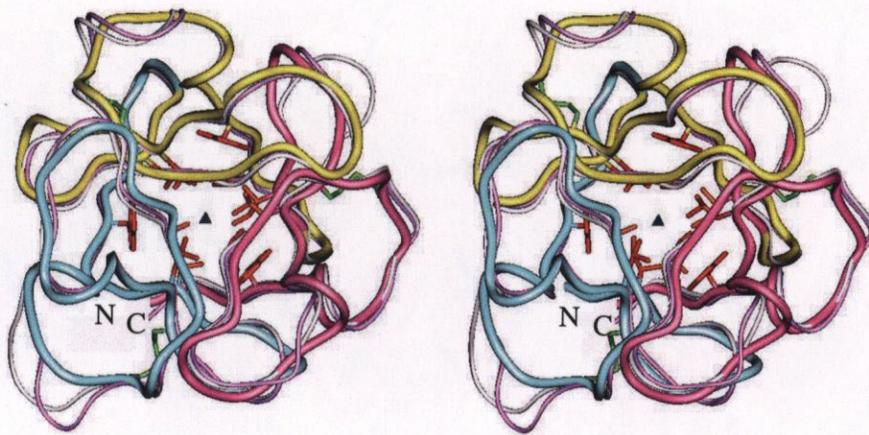


Fig. 2-7 Superposition of XBD (subdomains α ; blue, β ; yellow and γ ; pink), and the equivalent domains of ricin B-chain 1 (white; 2aai; Rutenber & Robertus, 1991) and abrin B-chain 1 (purple; 1abr; Tahirov *et al.*, 1995). Hydrophobic residues that form the hydrophobic core are shown in orange, and disulfide bonds in green.

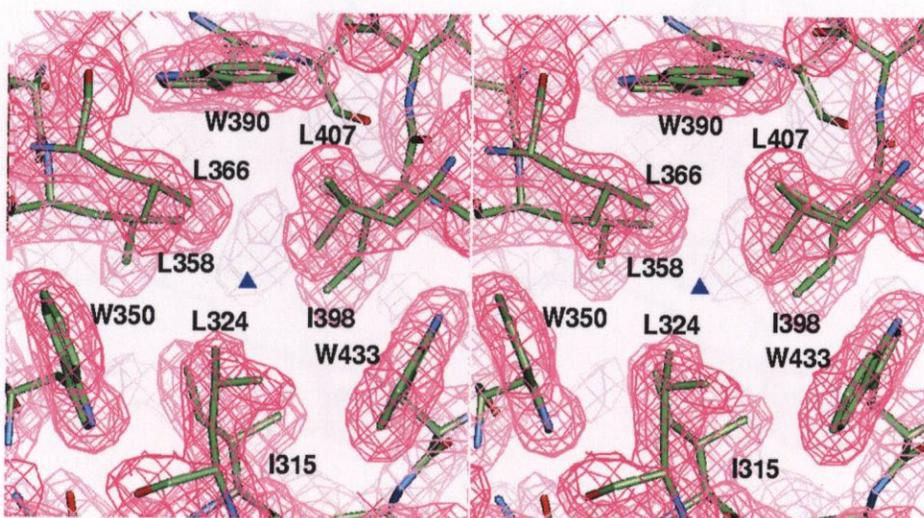


Fig. 2-8 Stereoview of the $2F_{\text{obs}} - F_{\text{calc}}$ electron density map around the hydrophobic core of XBD contoured at 1σ . Side-chains of Leu and Trp extend toward a pseudo-three-fold axis marked by a solid triangle, forming the top layer of the core cylinder. Similarly, side-chains of Ile (Leu in subdomain β) form the bottom layer of the core.

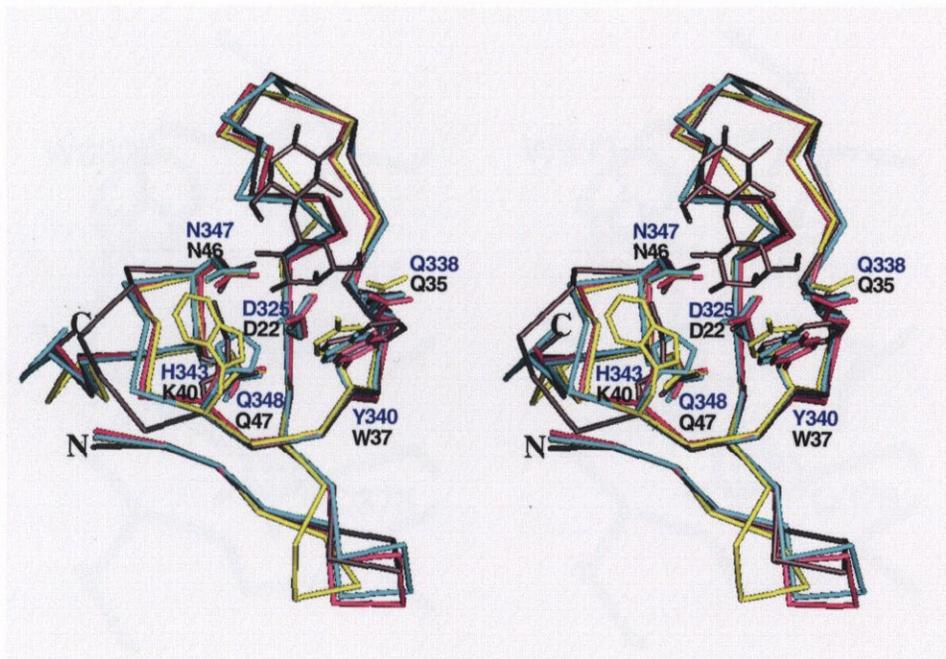


Fig. 2-9 Stereoview of the superposition of three subdomains of XBD in comparison with subdomain 1 α of a ricin B-chain/lactose complex (Rutenber & Robertus, 1991). Subdomains α , β , γ and ricin B-chain 1 α are colored in blue, yellow, pink and gray, respectively. Residues involved in the lactose-binding in ricin and the corresponding residues in XBD are shown. Residues of ricin and subdomain β of XBD are labeled in black and blue, respectively. Four polar residues are conserved among them, and Trp in ricin is replaced by Tyr in XBD.

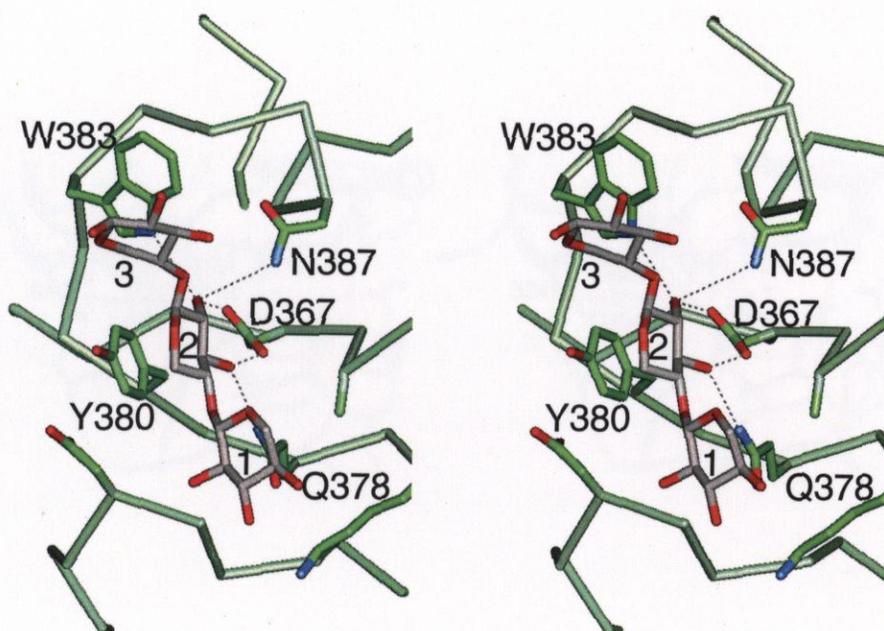


Fig. 2-10 Stereoview of xylotriose docking model in subdomain β of XBD. The xylotriose structure was built based on the X-ray structure of β -1,4-xylan hydrate (Nieduszynski & Marchessault, 1972) and fitted into the binding site manually. Sugars are numbered from the non-reducing end. XBD residues interact with the 2nd xylose sugar through five hydrogen bonds (broken lines). The 2nd and 3rd sugar rings are placed over the aromatic rings of Tyr380 and Trp383.

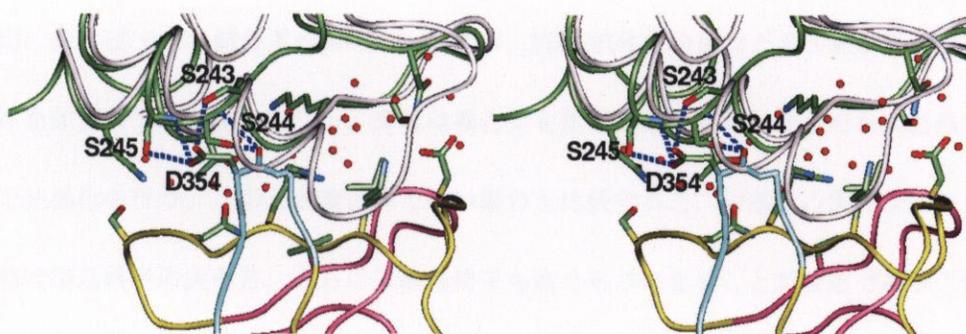


Fig. 2-11 Stereoview of an interacting region between the catalytic domain (green) and XBD (subdomains α ; blue, β ; yellow and γ ; pink) superimposed on the equivalent region of the catalytic domain in Cex (white; 2exo; White *et al.*, 1994). Five inferred hydrogen-bonding interactions between the triple Ser sequence in the N-terminal end of $C\alpha_7$ and Asp354 in XBD are shown as blue broken lines. The difference between FXYN and Cex is clearly seen in the region from $C\beta_7$ to $C\alpha_7$.