

**Molecular modelling studies**  
**with regard to the observed resistance profile**  
**of *Streptococcus pneumoniae* PBP2x**  
**resulting from some naturally occurring**  
**active site mutations**

**Inaugural-Dissertation**

Zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

**Paulus A. M. van Hooft**

aus Utrecht

Düsseldorf 2001







**Molecular modelling studies**  
**with regard to the observed resistance profile**  
**of *Streptococcus pneumoniae* PBP2x**  
**resulting from some naturally occurring**  
**active site mutations**

**Inaugural-Dissertation**

Zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

**Paulus A. M. van Hooft**

aus Utrecht

Düsseldorf 2001

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät  
der Heinrich-Heine-Universität Düsseldorf

1. Gutachter: Prof. Dr. H.-D. Höltje

2. Gutachter: Prof. Dr. P. Proksch

Tag der mündlichen Prüfung: 1. Dezember 2000

Für meine Eltern, für Julia und Laurens





Die vorliegende Arbeit wurde von Juli 1996 bis Juli 2000 am Institut für Pharmazeutische Chemie der Heinrich-Heine-Universität Düsseldorf auf Anregung und unter Anleitung von Herrn Prof. Dr. H.-D. Höltje angefertigt.

Mein besonderer Dank gilt Herrn Prof. Dr. H.-D. Höltje für die Überlassung des interessanten Themas, das gute Arbeitsklima und die hervorragenden Arbeitsbedingungen. Durch seine ständige Gesprächsbereitschaft und seine intensive Unterstützung hat er wesentlich zum Gelingen dieser Arbeit beigetragen. Ausserdem habe ich sein Interesse für die niederländische Kultur und Literatur immer geschätzt.

Herrn Prof. Dr. J.-M. Frère, Centre for Protein Engineering, Universität Lüttich, möchte ich ganz herzlich für die Einführung in die Kinetik der Penicillin-bindenden Proteine und  $\beta$ -Lactamasen danken.

Herrn Dr. O. Dideberg, Institut de Biologie Structurale, Grenoble, danke ich für die Überprüfung des Protein Modells mit den Kristallstrukturdaten.

Bei Frau Prof. Dr. R. Hakenbeck, Universität Kaiserslautern, Department of Microbiology, möchte ich mich besonders herzlich für die gute Zusammenarbeit bedanken.

Herrn Prof. Dr. P. Proksch danke ich herzlich für die Anfertigung des Zweitgutachtens.

Weiterhin möchte ich mich ganz besonders bei allen Kolleginnen und Kollegen des Arbeitskreises für die ausgesprochen freundschaftliche und konstruktive Zusammenarbeit und Hilfsbereitschaft, sowie auch für die ständige Diskussionsbereitschaft bedanken.



---

<b>INTRODUCTION .....</b>	<b>1</b>
1.1 The bacterial cell wall.....	4
1.2 Murein biosynthesis .....	6
1.2.1 Overview.....	6
1.2.2 The final step in cell wall biosynthesis .....	7
1.3 Penicillin-binding proteins .....	10
1.3.1 Protein family .....	10
1.3.2 Kinetics of the transpeptidase reaction catalysed by PBPs .....	13
1.4 Inhibition of penicillin-binding proteins .....	14
1.4.1 Penicillins and cephalosporins.....	14
1.4.2 Kinetics of acylation by $\beta$ -lactam antibiotics.....	16
1.4.3 Reaction with acyclic ester/thiolester carbonyl donors.....	17
1.5 Resistance to $\beta$ -lactam containing antibiotics .....	18
1.5.1 $\beta$ -Lactamase-mediated resistance.....	18
1.5.1.1 Mechanistic properties .....	18
1.5.1.2 Classification of $\beta$ -lactamases.....	20
1.5.1.3 Structural and mechanistic properties of active site serine proteins..	21
1.5.2 Intrinsic resistance .....	22
1.6 Resistance to $\beta$ -lactam antibiotics in <i>Streptococcus pneumoniae</i> .....	24
1.6.1 Introduction.....	24
1.6.2 Structure of <i>Streptococcus pneumoniae</i> PBP2x.....	25
1.6.3 Amino acid mutations of PBP2x.....	27
<b>2 AIM AND SCOPE OF THIS THESIS .....</b>	<b>31</b>
<b>3 METHODOLOGY.....</b>	<b>35</b>
3.1 Describing molecular systems .....	37
3.1.1 Fundamentals.....	37
3.1.2 Quantum chemical methods .....	39
3.1.2.1 <i>Ab initio</i> SCF MO methods .....	40
3.1.2.2 Semiempirical SCF MO methods.....	42
3.1.3 Force field methods .....	43

---

3.1.3.1	Backgrounds .....	43
3.1.3.2	Energy minimisation with force field methods .....	45
3.1.3.3	The Consistent Valence Force Field .....	46
3.1.4	Charges .....	50
3.2	Exploring the molecular configurational space.....	52
3.2.1	Macroscopic versus microscopic properties .....	52
3.2.2	Metropolis or Monte Carlo simulations.....	53
3.2.3	Molecular dynamics simulations .....	54
3.3	Comparative protein modelling .....	56
3.3.1	Introduction.....	56
3.3.2	Alignment methods.....	58
3.3.2.1	Background.....	58
3.3.2.2	Needleman-Wunsch algorithm for global alignments.....	59
3.3.2.3	Smith-Waterman algorithm for local alignments.....	60
3.3.2.4	Scoring functions for amino acid mutations .....	61
3.3.2.5	Gap penalty functions .....	61
3.3.3	Rigid-body assembly approach .....	62
3.3.4	Segment matching methods .....	62
3.3.5	Satisfaction of spatial restraints .....	63
3.3.6	Protein modelling starting from C <sup>α</sup> coordinates.....	63
3.3.6.1	Overview .....	63
3.3.6.2	The virtual-bond chain method.....	64
3.3.6.3	The dipole-path method .....	65
3.4	Molecular docking methods .....	66
3.4.1	Backgrounds.....	66
3.4.2	Rigid-body approach.....	67
3.4.3	Flexible-ligand docking .....	67
3.4.3.1	Introduction .....	67
3.4.3.2	AutoDock 2.4 .....	68
<b>4</b>	<b>RESULTS.....</b>	<b>73</b>
4.1	Generation of all-atom coordinates.....	75
4.2	Relaxation of the transpeptidase domain.....	81

---

4.2.1	General comments .....	81
4.2.2	Relaxation strategy .....	81
4.2.3	Relaxation of the backbone .....	83
4.2.4	Relaxation of the full transpeptidase domain .....	85
4.3	The final model of <i>S. pneumoniae</i> PBP2x .....	91
4.4	Inhibitorss .....	96
4.4.1	Choice of inhibitors .....	96
4.4.2	Construction of the ligands .....	97
4.5	Enzyme-inhibitor complexes .....	99
4.5.1	Docking inhibitors into the active site of PBP2x.....	99
4.5.2	Molecular dynamics simulations of PBP2x <sub>R6</sub> complexed with inhibitors	100
4.5.2.1	Simulation conditions .....	100
4.5.2.2	Simulation of PBP2x* <sub>R6</sub> complexed with benzylpenicillin .....	102
4.5.2.3	Simulation of PBP2x* <sub>R6</sub> complexed with cefotaxime .....	108
4.5.2.4	Simulation of PBP2x* <sub>R6</sub> complexed with oxacillin .....	114
4.5.2.5	Simulation of PBP2x* <sub>R6</sub> complexed with cephalothin .....	120
4.5.3	Molecular dynamics simulations of PBP2x* <sub>T526 → S</sub> complexed with inhibitors .....	127
4.5.3.1	Simulation of PBP2x* <sub>T526 → S</sub> complexed with benzylpenicillin .....	127
4.5.3.2	Simulation of PBP2x* <sub>T526 → S</sub> complexed with cefotaxime .....	134
4.5.4	Summary of the various simulations .....	140
<b>5</b>	<b>DISCUSSION .....</b>	<b>143</b>
5.1	Introductory aspects .....	145
5.2	The procedure for generating full structure from C <sup>α</sup> information .....	146
5.3	Generation of the transpeptidase domain of <i>S. pneumoniae</i> PBP2x* .....	148
5.3.1	Local structural trace alignment .....	148
5.3.2	Relaxation of the backbone .....	148
5.3.3	Relaxation of the whole transpeptidase domain .....	149
5.4	Homology model of the transpeptidase domain of PBP2x* .....	150
5.5	Complexes of PBP2x* <sub>R6</sub> .....	152
5.5.1	General remarks .....	152
5.5.2	PBP2x* <sub>R6</sub> complexed with benzylpenicillin and cefotaxime .....	152

---

5.5.3	PBP2x* <sub>R6</sub> complexed with oxacillin and cephalothin.....	154
5.5.4	Comparison of the complexes with PBP2x* <sub>R6</sub> .....	155
5.6	Complexes with PBP2x* <sub>T526→S</sub> .....	157
5.6.1	Introduction .....	157
5.6.2	Comparison of the complexes with PBP2x* <sub>T526→S</sub> .....	157
<b>6</b>	<b>SUMMARY / ZUSAMMENFASSUNG .....</b>	<b>159</b>
	<b>REFERENCES.....</b>	<b>163</b>
	<b>APPENDICES.....</b>	<b>179</b>
	Appendix A List of abbreviations.....	181
	Appendix B Amino acids .....	182
	Appendix C Hardware.....	183
	Appendix D Software .....	184
	Appendix E Additional CVFF parameters .....	185
	Appendix F Bcl macros.....	187
	Appendix G Awk scripts.....	196
	Appendix H Analysis of the molecular dynamics simulations.....	197
	Appendix I Docking scripts for AutoDock 2.4 .....	200
	Appendix J Input files for DISCOVER .....	202

## Mathematical constants

e	2.71828
$\pi$	3.14159

## Fundamental constants and units

e	elementary charge of one electron ( $1.60219 \times 10^{-19}$ C)
$\epsilon_0$	vacuum permittivity ( $8.85419 \times 10^{-12}$ C <sup>2</sup> ·J <sup>-1</sup> ·m <sup>-1</sup> )
h	Planck constant ( $6.62618 \times 10^{-34}$ J·s)
$\hbar$	Dirac constant, or $h/2\pi$ ( $1.05459 \times 10^{-34}$ J·s)
$k_B$	Boltzmann constant ( $1.38066 \times 10^{-23}$ J·K <sup>-1</sup> )
$m_u$	atomic mass unit ( $1.66056 \times 10^{-27}$ kg) (equals the mass of one hydrogen atom, or one Dalton (Da))
$N_A$	Avogadro constant ( $6.02205 \times 10^{23}$ mol <sup>-1</sup> )
R	Gas constant ( $8.31441 \times$ J·K <sup>-1</sup> ·mol <sup>-1</sup> )

## Useful relations

Å	Ångstrom ( $10^{-10}$ m)
KDa	kiloDalton ( $10^3$ Da)
fs	femtosecond ( $10^{-15}$ s)
ps	picosecond ( $10^{-12}$ s)





## **1 Introduction**

---



One of the most important developments in health care in the twentieth century originated from the observation by Alexander Fleming in 1928 [1], that bacterial growth was inhibited by a contaminating mould (*Penicillium*). He was encouraged by his findings that it was not toxic to animals after injection of an extract of the mould. Despite efforts to isolate the antibacterial substance himself, it was not until the early nineteen-forties, before penicillin could be isolated by Howard Florey and Ernest Chain [2], which has led to an effective, widespread clinical use of this antibiotic. Since then, many derivatives have been successfully used to fight bacterial diseases. However, the mechanism of action of these compounds did not become clear until the 1960s, when most of the cell wall components have been identified. Especially the surprising discovery by Murray et al. [3] that the major component of the cell wall in the highly susceptible Gram-positive bacteria (peptidoglycan or murein) is also present in the much less susceptible Gram-negative species, has led to a vivid picture about the basic murein structure in bacteria. At that time the mechanism of action of antibiotics was believed to be completely understood [4]. A few years later, however, it was observed that the presence of antibiotics did not have to result in lysis per se, but could also cause the bacteria to just stop growing (bacteriostasis). Thus, it was soon recognised that the mechanism of action of antibiotics was much more complicated than previously thought. Especially during the last fifteen years, new insights into the complex nature about the dynamics, function and fine molecular structure have been reached and are still developing.

As a natural reaction to the lethal effect of these compounds, however, bacteria have developed strategies, which enable them to survive in the presence of these antibiotics. The most important of these strategies is the development of defensive proteins ( $\beta$ -lactamases), the action of which is the denaturation of these antibiotics into inactive compounds. Other mechanisms include reduction of the permeability of the cell wall for these compounds and modification of the target enzymes of these lethal compounds (penicillin-binding proteins). An example of a species adopting this strategy in order to survive in the presence of antibiotics is *Streptococcus pneumoniae*, an important pathogenic agent in man. Although highly alarming, its developing resistance is quite an astonishing achievement, since these modified proteins have still to be able to perform their natural task(s).

## 1.1 The bacterial cell wall

In the broadest sense, antibiotics are natural substances, which display their action by eliminating self-replicable and illness-causing agents out of mammals. This definition includes metabolites of both microorganisms and higher organisms acting against either tumour cells or viruses. Because of differences in structure and biosynthesis of both the cell wall and ribosomes in procaryotic and eucaryotic organisms, these cellular sites form important targets for antibiotics.

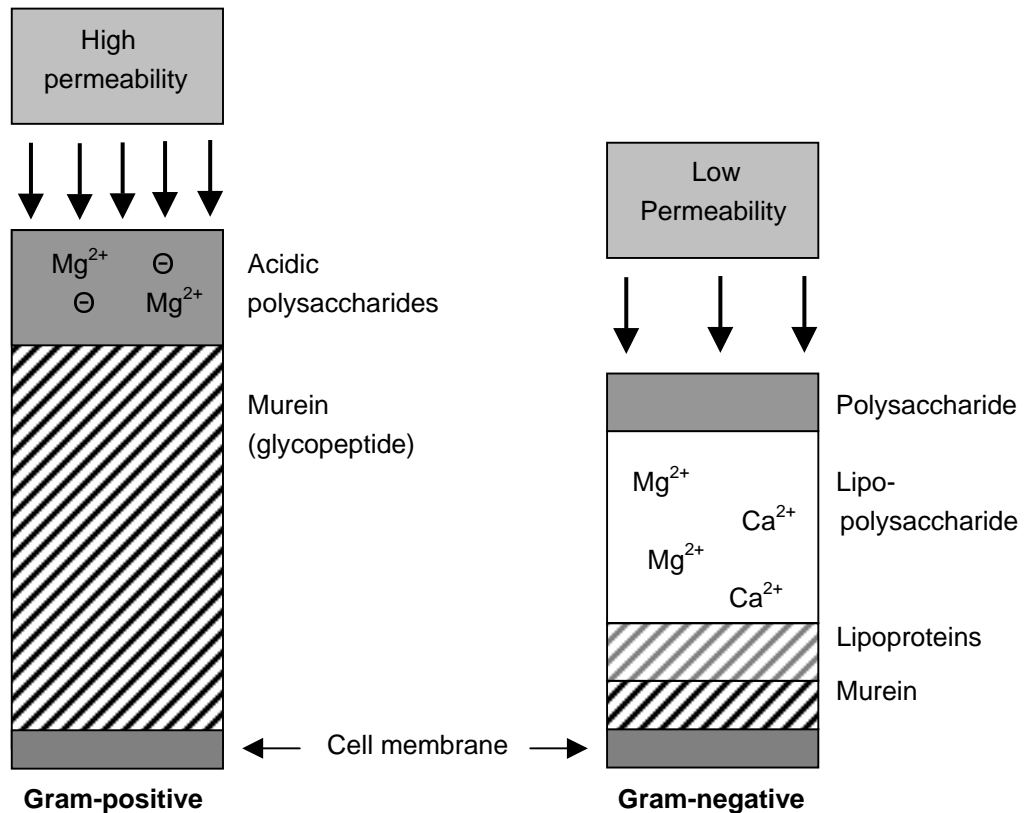
In order to reach these cellular sites, antibiotics have to be transported into the cell either actively or passively. In the case of active transport the antibiotic is passed through transport systems (porins), which are located in the plasmamembrane. Highest selectivity is achieved here if transport is carried out by systems that do not occur in mammalian cells. With passive transport on the other hand, the antibiotic has to cross the cell wall and in most cases the plasma membrane as well.

The cell wall of Gram-positive bacteria is composed of one thick murein layer (peptidoglycan), onto which acidic polysaccharides (teichuroic acid and teichuronic acid) are covalently connected (Figure 1.1). It functions as a molecular sieve, but does not represent a permeability barrier to most antibiotics. On the other hand, the cell wall of Gram-negative bacteria is characterised by a lipid membrane, which is situated between the outer membrane and the cytoplasmic membrane. A single layer of murein is situated in the periplasm [5], which covers the space between the outer membrane and the cytoplasmic membrane. The murein layer is connected to the latter membrane by lipoproteins.

The asymmetric character of the lipid membrane mentioned above is responsible for the low permeability of the cell wall of Gram-negative bacteria. Whereas the inner side of this membrane consists of lipoproteins, the outer part (which is connected to the polysaccharide-containing outer membrane) consists of lipopolysaccharides. Because of the fact that the outer lipid membrane contains water-filled porins, it is only permeable for small hydrophilic molecules.

On the other hand, it is the peptidoglycan structure that provides the bacterial sacculus with a mechanical stability as to withstand the intracellular pressure [6, 7]. Although the chemical composition of the polymer may vary between different bacteria, the basic architecture is conserved. It is a heteropolymer consisting of

relatively short glycan strands, composed of saccharide subunits, which are crosslinked by peptides.



**Figure 1.1** Schematic representation of the difference in bacterial cell wall composition in gram-positive and gram-negative bacteria [8].

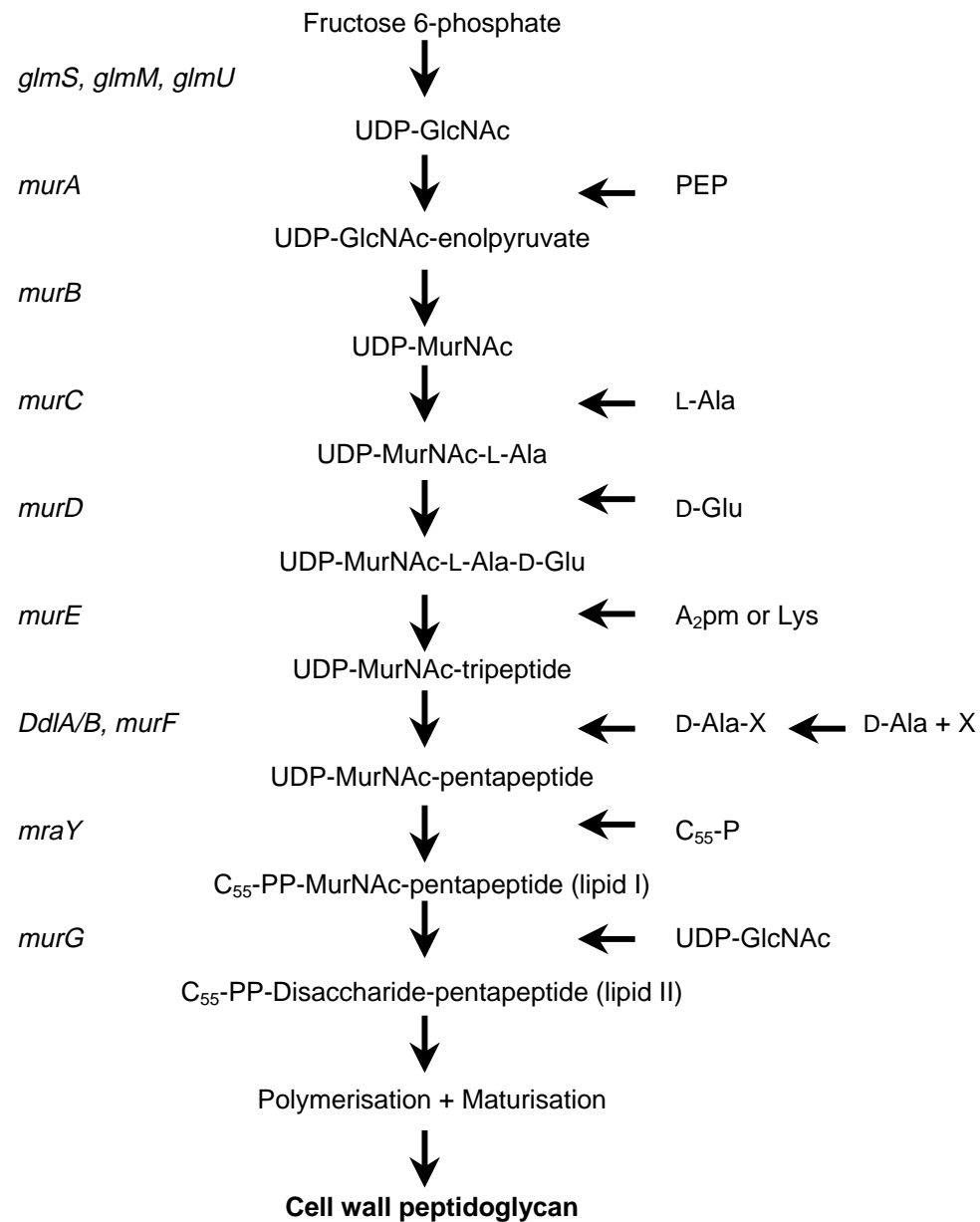
It has been shown that, in contrast to the chemically closely related chitin, peptidoglycan is not a crystalline structure [9]. Rather, its mechanical properties are more those of a viscous gel [10], the elastic properties of which reside within the conformational freedom of the peptide crosslink, that can be stretched up to fourfold [11]. Secondary structure can only be adopted by the sugar strands, which are far more helical.

## 1.2 Murein biosynthesis

### 1.2.1 Overview

The biosynthesis of peptidoglycan is a complex two-stage process [12, 13]. The first stage concerns the synthesis of its monomer unit, which is carried out by enzymes located in the cytoplasm or at the inner side of the cytoplasmic membrane. The final product of the first stage is the lipid intermediate undecaprenol. It has to be transported across the membrane, since it serves as the initial substrate for the polymerisation reactions taking place at the outer side of the cytoplasmic membrane during the second stage.

The final assembly of the monomer unit during the second stage can be divided into four groups of reactions: formations of uridine 5'-pyrophosphate-*N*-acetylglucosamine (UDP-GlcNAc), uridine 5'-pyrophosphate-*N*-acetylmuramic acid (UDP-MurNAc), UDP-MurNAc-peptides and lipid intermediates (Figure 1.1).

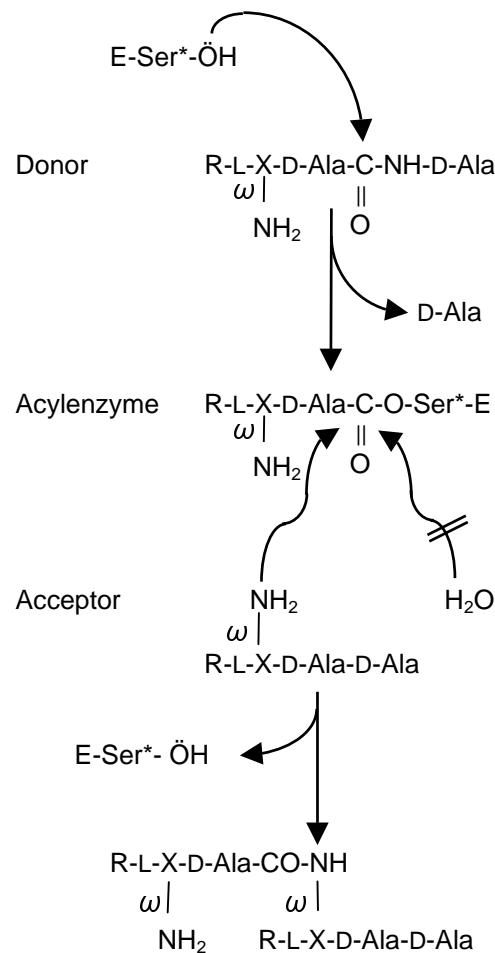


**Figure 1.1** *Assembly of the peptidoglycan monomer unit.*

### 1.2.2 The final step in cell wall biosynthesis

As mentioned on page 3 the murein structure present in the cell wall of bacteria represents the target for  $\beta$ -lactam-containing antibiotics. These agents are able to inhibit the machinery of the cell wall biosynthesis by inhibiting the final step of the cell wall biosynthesis, which takes place on the outer side of the cytoplasmic membrane

and is therefore readily accessible. As explained in Figure 1.1 the final step in murein biosynthesis consists of polymerisation of the C<sub>55</sub>-disaccharide-pentapeptide units (transglycosylation) and penicillin-sensitive formation of crosslinkages between the  $\omega$ -amino terminal of meso-A<sub>2</sub>pm (D-center) and the carboxyterminal of the penultimate D-alanine residue of the pentapeptide side chains (transpeptidation) [14, 15, 16]. For each cross-linkage formed, one molecule of the terminal D-alanine residue of the pentapeptide unit is removed. Alternatively, carboxypeptidation occurs when the accepting group is a water molecule, thereby limiting the extent of crosslinking [17].



**Figure 1.1** Schematic representation of the transpeptidase reaction. Ser\* represents the active serine.

Both transpeptidation and transglycosylation are carried out by unique bacterial proteins, called penicillin-binding proteins (PBPs, see section 1.3). They belong to the active site serine protein family. Because of the unique optical specificity that they



---

exhibit [18] (the scissile peptide bond extends between two D-alanine residues in  $\alpha$ -position to a free carboxylate group), they are also D,D-peptidases.

## 1.3 Penicillin-binding proteins

### 1.3.1 Protein family

PBPs were first separated [19] by column chromatography in 1972, but at that time it was not understood, why there were multiple penicillin-binding components. Three years later [20], six of these proteins (PBP1 to PBP6) were separated by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) and fluorography, which were subsequently grouped in the high molecular mass (high- $M_r$ , PBP1 to 3) and low molecular mass (low- $M_r$ , PBP4 to 6) PBPs. The probable functions of each of the PBPs of *Escherichia coli* were ascertained by isolating mutants defective in some of these proteins. It was found that the highest molecular weight PBP1 (separated afterwards into PBP1A and a group of PBPs, PBP1B- $\alpha$ , - $\beta$  and - $\gamma$ ) [21, 22, 23] was believed to function in cell elongation, PBP2 in determination of the cell-shape, and PBP3 in the formation of septa [24]. Today, eight kinds of PBPs have been identified (PBP1-8), most of which can be subdivided. Each bacterial species has at least three and up to more than eight PBPs [25], whose function *in vivo* is a matter of controversy, if not unknown.

An evolutionary scenario has been proposed [26], through which acquisition of new functions from a putative DD-transpeptidase/PBP ancestor is achieved by local changes. According to this scenario, the low- $M_r$  PBPs, which are able to catalyse hydrolysis of the ester bond of the peptidyl enzyme, have evolved from the high- $M_r$  PBPs. Thus, the low- $M_r$  PBPs, which indeed have a single catalytic function (they are mainly D,D-carboxypeptidases or peptidoglycan hydrolases [27]) are able to control the extent of cross-linking by the high- $M_r$  PBPs.

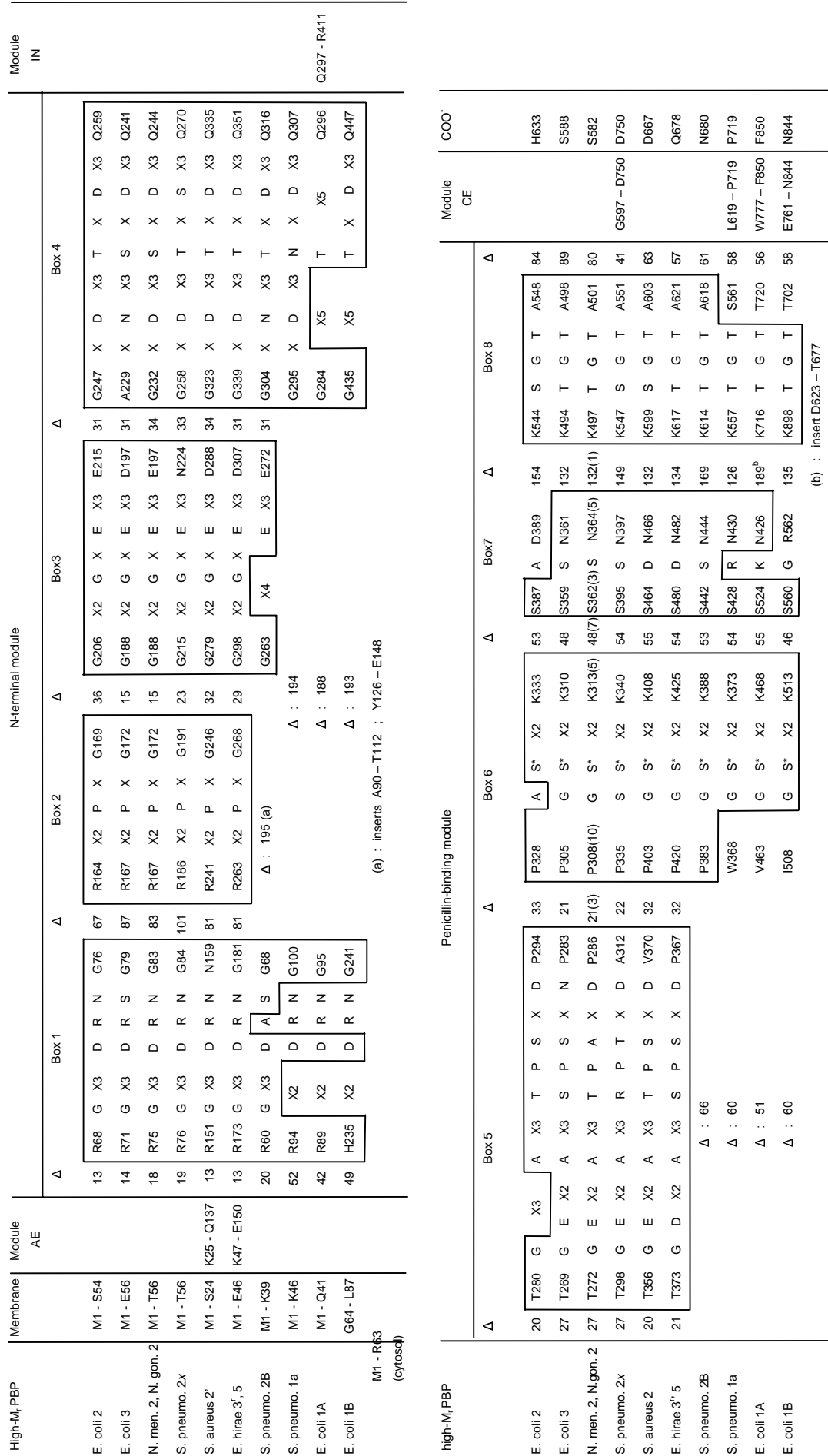
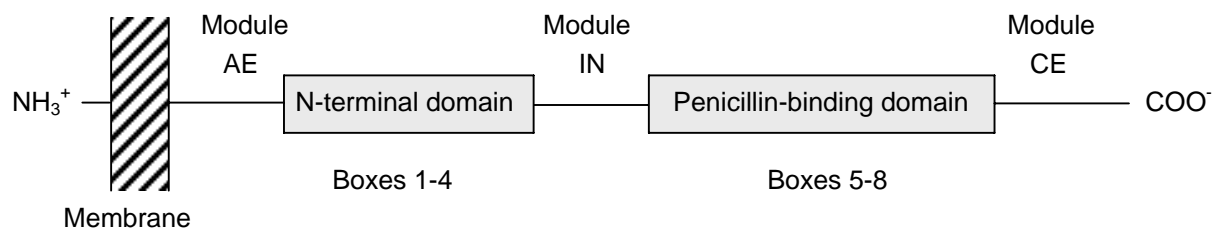


Figure 1.4 Modular design of the bi(multi)functional high molecular mass PBPs. See also Figure 1.5.

The high- $M_r$  PBPs are bi(multi)functional proteins [26] and may exhibit both transglycosylase and transpeptidase activities. They are essentially constructed of two modules, a penicillin-binding (PB) module and a non-penicillin-binding (n-PB) module, but show increasing complexity resulting from the acquisition of additional modules. These are linked together in a single polypeptide chain [28] that folds on the outer side of the plasma membrane and is anchored into the membrane by an uncleaved amino-terminal signal peptide. In some high- $M_r$  PBPs an amino extension 'AE' is inserted between the membrane anchor and the N-terminal module. An internal extension 'IN' may additionally be fused between the N-terminal module and the PB module, and a carboxy extension 'CE' may be fused to the carboxy end of the PB module (see Figures 1.4 and 1.5). Depending on the motifs present in the n-PB modules [28, 29, 30], high- $M_r$  PBPs belong to Class A or Class B.



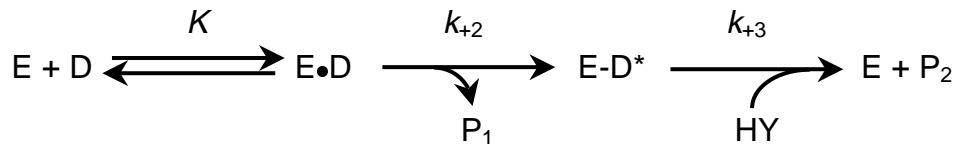
**Figure 1.5** Modular design of the bi(multifunctional high molecular mass PBPs. The box numbers refer to Figure 1.4.

In contrast to the low- $M_r$  PBPs, whose function is interchangeable, the high- $M_r$  PBPs of class A and B are thought to be indispensable. Upon inhibition of these proteins the cell wall is no longer able to withstand the internal pressure and cell lysis will result.

High molecular mass PBPs that have no function in cell wall peptidoglycan metabolism also exist, forming class C high- $M_r$  PBPs. Members discovered so far [31, 32] include the BLaR 601-amino-acid PBP of *Bacillus licheniformis*, *S. aureus* MECR1 and *S. epidermidis* MECR1. They function as penicillin-sensory transducers invoking  $\beta$ -lactamase synthesis (see Section 1.5).

### 1.3.2 Kinetics of the transpeptidase reaction catalysed by PBPs

If E is the enzyme, D the carbonyl-donor substrate, HY a nucleophilic acceptor and P<sub>1</sub> and P<sub>2</sub> reaction products the D,D-peptidases catalyse the following reaction [16]:



Scheme 1

where  $K$  is the dissociation constant of the Michaelis complex  $E \cdot D$  and  $k_{+2}$  and  $k_{+3}$  are first order rate constants with  $[D] \gg [E]$  and if it is assumed that  $E \cdot D$  is in rapid equilibrium with both the free enzyme and the carbonyl donor substrate. If, furthermore,  $[D] \gg K$ , the catalytic efficiency at low carbonyl donor substrate concentrations and  $k_{+2}/K$  is the second order rate constant of enzyme acylation, the catalytic efficiency of the reaction above is

$$\frac{k_{\text{cat}}}{K_m} = \frac{k_{+2}}{K} \quad (1.1)$$

The  $k_{\text{cat}}/K_m$  parameter is independent of the deacylation rate and corresponds to the second-order rate constant for the acylation reaction [16].

## 1.4 Inhibition of penicillin-binding proteins

Compounds that mimic the transition state of the transpeptidation reaction are able to inhibit PBPs. Thus [33, 34, 35],  $\beta$ -Lactam containing compounds such as penicillins, 6-epoxy-spiropenicillins, cephalosporins, carbapenems, monobactams and some other compounds like thienamycin, bicyclomycin, lactivicin and pyrazolidone are able to mimic the transition state of the D-alanine-D-alanine terminal of the donor peptidoglycan strand in the transpeptidation reaction. These antibiotics are therefore able to inhibit PBPs by forming a covalently bound serine ester-linked acyl enzyme.

### 1.4.1 Penicillins and cephalosporins

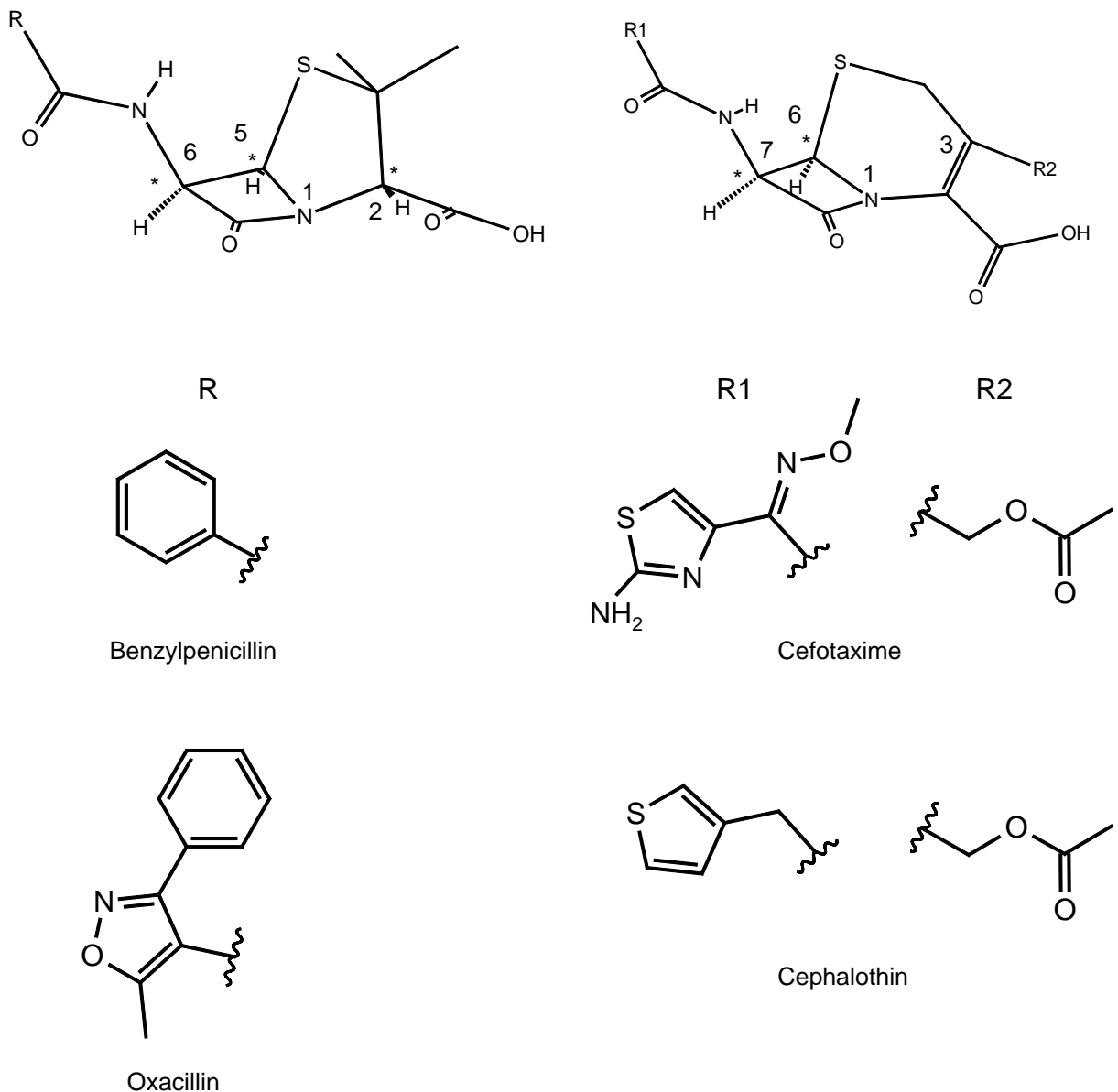
Penicillins and cephalosporins represent the most important classes of  $\beta$ -lactam antibiotics. As shown in Figure 1.6 the four-membered  $\beta$ -lactam ring in these compounds is part of a bicyclic ringsystem. In penicillins the fused ring system is a thiazolidine ring, whereas cephalosporins are characterised by a dihydrothiazine ring system. Penicillin itself (6-aminopenicilloic acid) is a cyclic dipeptide composed of D-valine and L-cysteine.

Based on the structure of the side chain (see also Section 1.5), penicillins can be classified in the following groups:

- orally active penicillins (phenoxyphenicillins)
- orally and parentally active broad-spectrum penicillins ( $\alpha$ -aminobenzylpenicillins)
  - ampicillin and derivatives
  - carboxypenicillins
- orally and parentally applicable lactamase-stable penicillins (isoxazolylpenicillins)

Likewise, cephalosporins can be classified as follows:

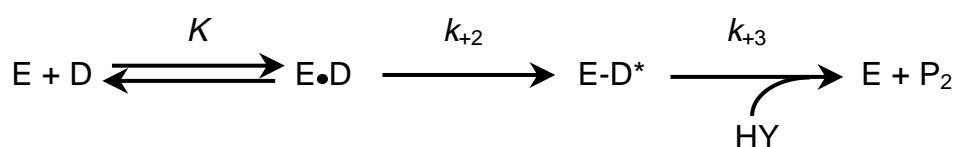
- orally or parenterally applicable (first generation) cephalosporins
- parenterally applicable lactamase-stable cephalosporins
- parenterally applicable (first generation) cephalosporins
- orally applicable lactamase-stable (third generation) cephalosporins
- parenterally active cephalosporinase-stable (fourth-generation) cephalosporin



**Figure 1.6** Basic structures of penicillins and cephalosporins, including two examples of each class of antibiotics.

### 1.4.2 Kinetics of acylation by $\beta$ -lactam antibiotics

Formation of a serine ester-linked acylenzyme is achieved upon nucleophilic attack of the  $\gamma$ -O of the active serine on the carbonyl-carbon, with concomitant bond cleaving of the amide bond of the  $\beta$ -lactam ring. The mechanism is similar to that of catalysed rupture of the peptide bond by the serine peptidases of the trypsin and subtilisin families [27]. Although it has never been demonstrated, it is well accepted that PBPs interact with  $\beta$ -lactam antibiotics according to the following three-step mechanism (compare with Figure 1.7 on page 18) [16, 36]:



Scheme 2

Apart from the fact that as a consequence of the endocyclic nature of the scissile amide bond in the  $\beta$ -lactam ring the leaving group  $P_1$  remains part of the acyl-enzyme, this reaction scheme is the same as that of scheme 1. Moreover,  $k_{+3}$  has a small absolute value.

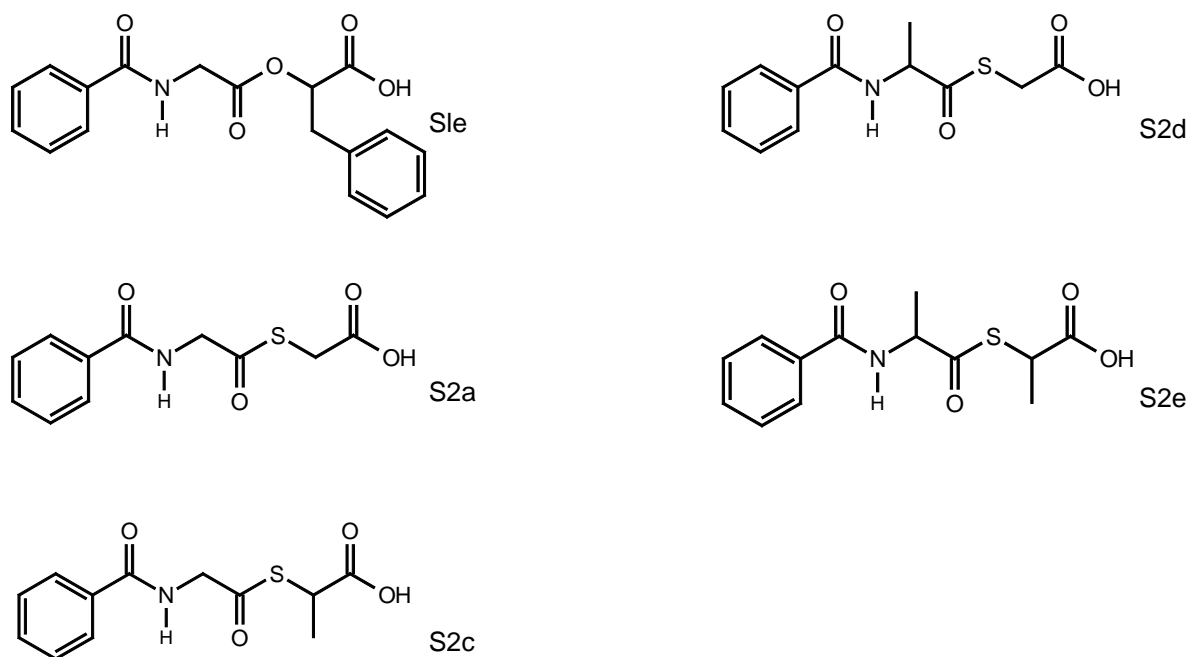
First, a Michaelis-Menten complex is formed between the enzyme and the donor molecule, leading to formation of an acyl(R-D-alanyl)-enzyme ( $E-D^*$ ). Upon nucleophilic attack of a water molecule or an amino-compound the enzyme is regenerated and the product is eliminated. However, in the case of PBPs the last step is (very) slow. Since only very simple amino nucleophiles, such as hydroxylamine, are able to attack the benzylpenicilloyl-enzyme [37], the benzylpenicilloyl moiety is likely to prevent access to the amino-acceptor site, resulting in complete inertness of the enzyme.



### 1.4.3 Reaction with acyclic ester/thioester carbonyl donors

Frère et al. [38, 39, 40] have shown that ester and thioester peptides (acyclic analogues of penicillin) are hydrolysed by all the penicilloyl serine transferases (including the  $\beta$ -lactamases, see section 1.6.1). This enables facile spectroscopic study of possible transpeptidation activities of high- $M_r$  PBPs, since both the  $k_{\text{cat}}$  and  $K_m$  parameters can be determined [40] by recording the initial rates at different substrate concentrations. In addition, the  $k_3$  parameter can be determined by spectroscopic measurement of the recovery of the thioesterase activity of these enzymes. Some frequently used ester peptides are shown in Table 1.1.

**Table 1.1** Structures of some frequently used ester peptide analogues for studying the catalytic properties of penicilloyl serine transferases.



These substances also greatly facilitate the study of effects of amino acid mutations on the catalytic profile of such enzymes as obtained by protein engineering or discovered in penicillin-resistant strains.

## 1.5 Resistance to $\beta$ -lactam containing antibiotics

During the last 50 years, however, bacteria have developed several strategies in order to be able to survive in the presence of these  $\beta$ -lactam antibiotics. Mechanisms resulting in resistance can be either genetic (both chromosome coded and plasmid coded) or biochemical. These strategies can be divided into [41, 42]:

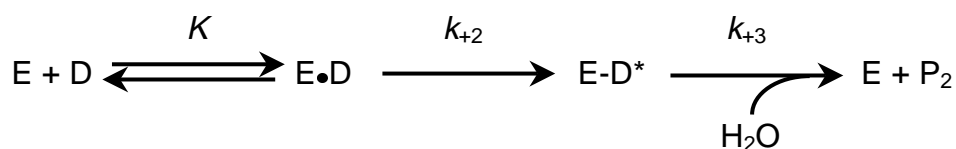
- (a) Synthesis of enzymes that inactivate  $\beta$ -lactam antibiotics (acylases, esterases or  $\beta$ -lactamases)
- (b) Reduction of the affinity for  $\beta$ -lactam antibiotics by modification of PBPs (intrinsic resistance)
- (c) Reduction of drug accumulation at the action site by modification of the permeation barrier

### 1.5.1 $\beta$ -Lactamase-mediated resistance

By far the most important defence mechanism is the production of  $\beta$ -lactamases, small proteins which are able to quickly denature  $\beta$ -lactam containing compounds. They form the third group of the family of penicillin-interactive proteins and the majority of them are also acyl serine transferases. To date, over 190 different  $\beta$ -lactamases have been identified [43].

#### 1.5.1.1 Mechanistic properties

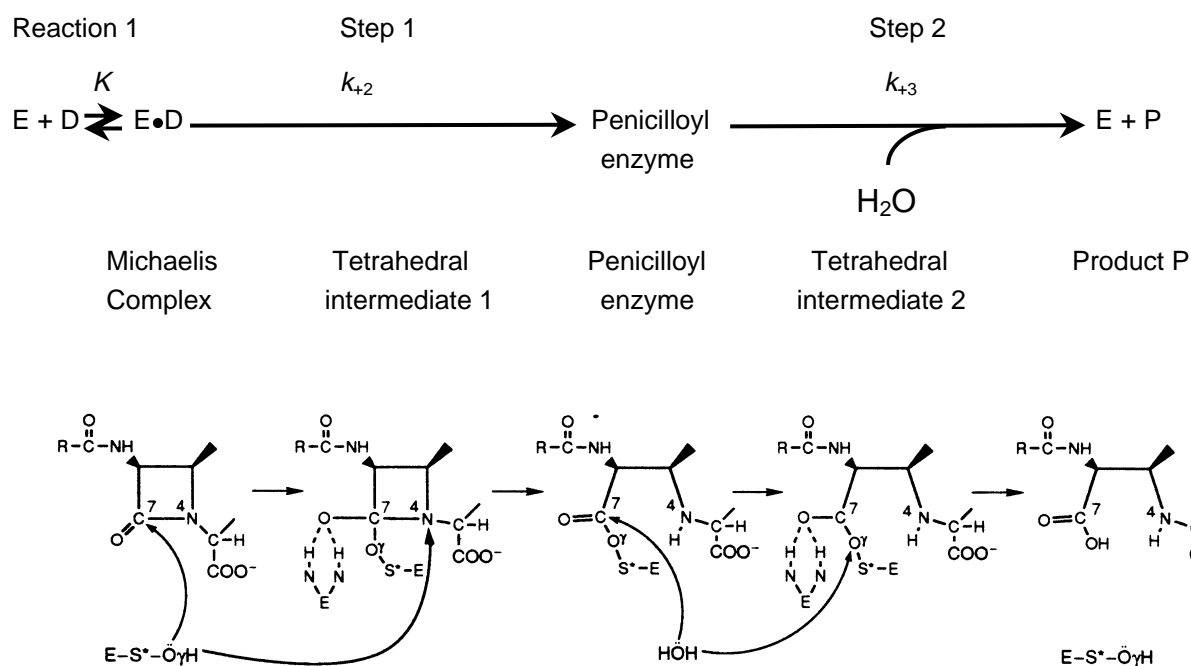
$\beta$ -Lactamases, however, do not recognise the donor peptide substrates of the peptidases, and the acylenzyme is effectively hydrolysed instead. The donor in reaction scheme 2 is now a water molecule:



Scheme 3

with both  $k_{+2}/K$  and  $k_{+3}$  very high. So, whereas  $\beta$ -lactamases catalyse hydrolysis of  $\beta$ -lactam antibiotics but cannot catalyse the hydrolysis of acyclic peptides, the DD-peptidases cannot efficiently catalyse  $\beta$ -lactam hydrolysis.

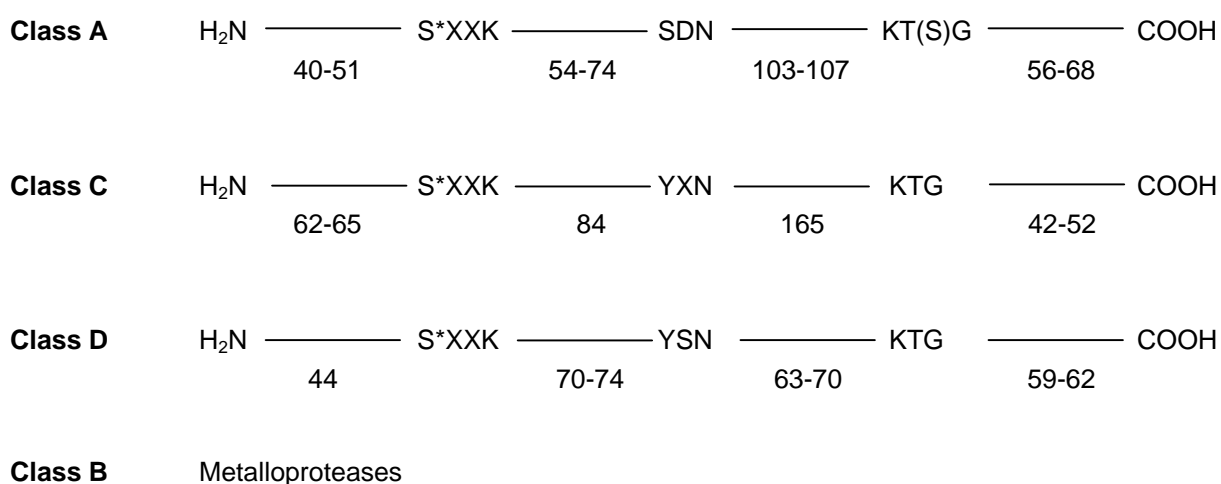
Both acylation and deacylation are regulated via an efficient proton abstraction-donation mechanism [44], which is carried out by other amino acids in the active site and structural water molecules (see Figure 1.7). There is much controversy about this mechanism, though.



**Figure 1.7** Proposed mechanism of proton abstraction-donation in  $\beta$ -lactam hydrolysing  $\beta$ -lactamases [44].

### 1.5.1.2 Classification of $\beta$ -lactamases

Based on the nucleotide sequence of their coding genes and the resulting amino acid sequence, Ambler [45] grouped the  $\beta$ -lactamases into classes A, B, C and D:



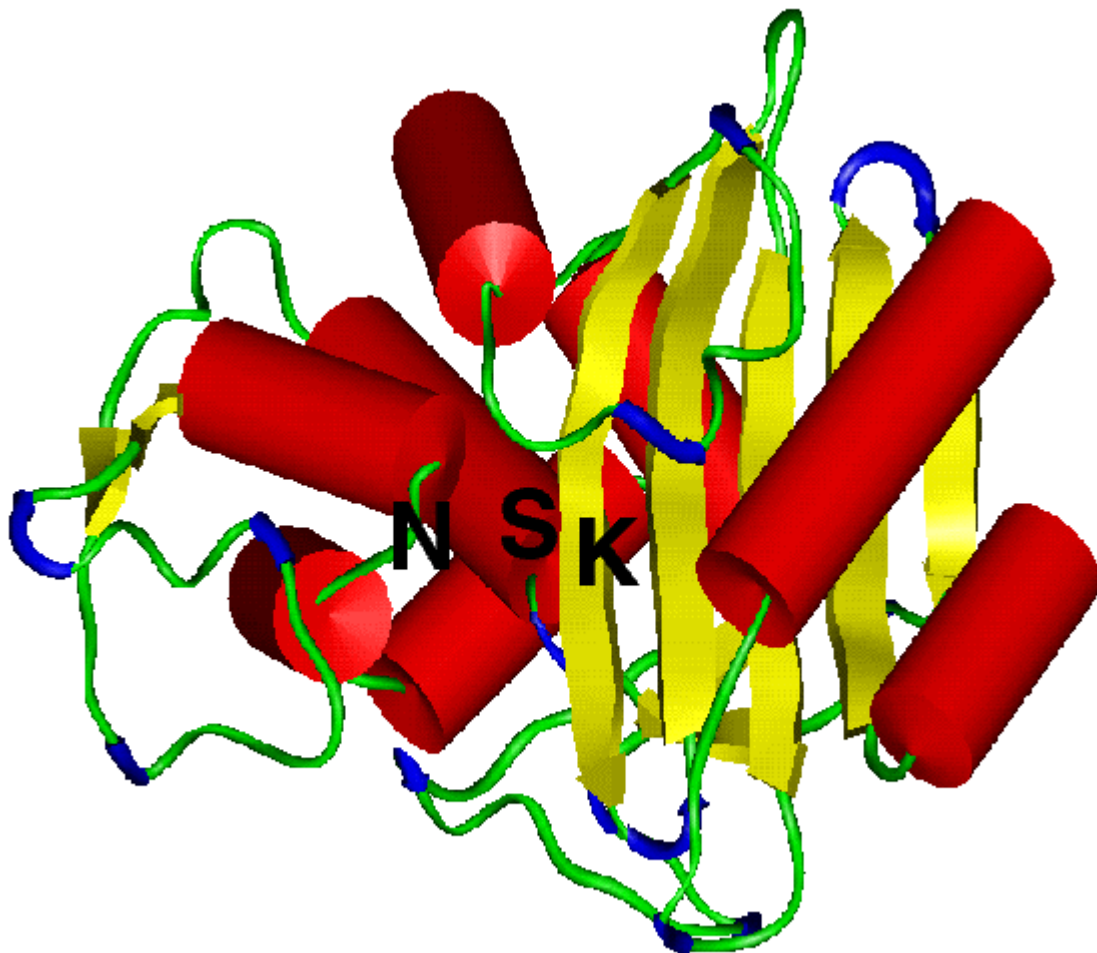
**Figure 1.8** Classification of  $\beta$ -lactamases according to their primary structure [45].

Only  $\beta$ -lactamases belonging to Class A, C and D are active-site serine acyl transferases or penicilloyl serine transferases. Members of Class B contain a zinc ion in their active site. As shown in Figure 1.8, members of the first three classes contain three structural motifs in their amino acid sequence: one tetrad S\*XXK containing the active site serine and two triads, S(Y)XN and K(H)T(S)G. Through folding of the peptide sequence, these three motifs are brought in close proximity to each other, forming the active site. In addition, between the second and third motif class A  $\beta$ -lactamases possess a conserved pentapeptide E166XELN170 situated on a so-called ‘omega’ loop, which has been shown [46] to take part in the proton abstraction-donation mechanism by keeping a conserved water molecule close in the vicinity of the active site.

Based on the catalytic profiles of these enzymes [47], class A enzymes are generally recognised as penicillinases, class C enzymes as cephalosporinases and class D enzymes as oxacillinases.

### 1.5.1.3 Structural and mechanistic properties of active site serine proteins

Although not all available, three-dimensional structures of eight class A  $\beta$ -lactamases are known [48]. In each of them the active site is located at the junction of an all- $\alpha$  and an  $\alpha$ - $\beta$  domain, the core of which is a five-stranded  $\beta$ -sheet, that is stabilised by  $\alpha$ -helices on both sides (see Figure 1.9). The tetrad S\*XXK is located at the amino terminal of an  $\alpha$ -helix, the triad S(Y)XN on a loop connecting two helices on one side of the active site cavity, and opposite to the S(Y)XN motif the tetrad K(H)T(S)G on the innermost strand ( $\beta_3$ ) of the five-stranded  $\beta$ -sheet.



**Figure 1.9** *Three-dimensional structure of Bacillus licheniformis 749/C  $\beta$ -lactamase. The structure is rendered according to Kabsch and Sander. The location of the active site is indicated by S (S\*XXK), N (S(Y)XN) and K (K(H)T(S)G).*

Although amino acid sequence homology between the various classes of  $\beta$ -lactamases and PBPs is rather low [49, 50], it has been demonstrated that this arrangement of secondary structure elements and motifs is also contained in the functionally different low- $M_r$  PBPs and the transpeptidase domain of the high- $M_r$  PBPs.

Upon binding, the negatively charged carboxylate group of penicillins and cephalosporins is electrostatically attracted [51] by the positively charged structurally preserved lysine residues in the active site. In the Michaelis-complex formed, a salt-bridge is formed between this carboxylate group and the positively charged  $N_\epsilon$  of the lysine contained in the K(H)T(S)G triad. Concomitantly, the  $\beta$ -lactam carbonyl oxygen is highly stabilised in the oxyanion hole formed by the backbone NH-groups of the active serine itself and the amino acid just below the K(H)T(S)G motif. As a result, the lactam carbonyl group becomes highly polarised. Upon nucleophilic attack of the  $\gamma O$  of the active serine on the carbonyl carbon atom the lactam ring is opened, resulting in ester bond formation with the enzyme. Ring opening is also driven by release of strain present in the four-membered lactam ring.

### 1.5.2 Intrinsic resistance

However, structural modification of cell components (points b and c mentioned on page 18) can be highly effective as well. For example, *Streptococcus* species such as *S. pneumoniae*, *S. mitis*, *S. oralis* and *S. sanguis* and *Neisseria* species are able to achieve resistance to  $\beta$ -lactam antibiotics by substituting some amino acid residues of their PBPs (see also section 1.6.3). This phenomenon, which was first reported in the 1970s [52], emerges via intra- and interspecies transfer of genes (mosaic genes) coding for their PBPs [53, 54, 55]. The goal of these mutations is to reduce the affinity of these proteins to  $\beta$ -lactams rather than affecting the deacylation step.

Another example of intrinsic resistance is the production of *additional*, low-affinity PBPs, as is the case in methicillin-resistant *Staphylococcus aureus* [56]. In this species, resistance is caused by the production of PBP2a or PBP2', which are only synthesised after the normal PBPs have been inactivated. By leaving the normal

---

PBPs unaltered, however, *S. aureus* does not face the same problem as *S. pneumoniae*, namely how to modify the PBPs so that the affinity towards  $\beta$ -lactam antibiotics is decreased, while still being able to perform their natural task.

## 1.6 Resistance to $\beta$ -lactam antibiotics in *Streptococcus pneumoniae*

### 1.6.1 Introduction

The genus *Streptococcus* (derived from the Greek streptos, twisted; kokkos, a bunch of grapes) is a genus of non-motile, non-spore forming, aerobic to facultatively anaerobic bacteria. These Gram positive bacteria occur in pairs or chains. The genus comprises many species, which occur regularly in the mouth and intestines of humans or animals. Some species are pathogenic and *Streptococcus pneumoniae* (or pneumococci) is the cause of considerable mortality and morbidity in humans. Pneumococci were the subject of pioneering genetic research work by Griffith when, in 1928, he first demonstrated genetic transformation in *S. pneumoniae*.

As mentioned in section 1.5.2, resistance in *Streptococcus pneumoniae* species is indeed mediated by structural modification of PBPs. This species contains six PBPs [57]. Five of them (PBP1a, 1b, 2x, 2a and 2b) are high- $M_r$  PBPs, whereas the sixth (PBP3) is a low- $M_r$  PBP. It might therefore be expected (see section 1.3.1), that PBP3, which acts as a DD-carboxypeptidase [58], is not essential. However, inactivation of PBP3 results in growth abnormalities [59]. It has been shown [60, 61, 62], though, that PBP2x and PBP2b (both belonging to Class B [30]) in *S. pneumoniae* are primary resistance determinants, since inhibition of PBP2x causes the cells to stop growing, whereas inhibition of PBP2b results in cell lysis. The essentiality of these two PBPs has further been demonstrated [63], in the sense that transformants lacking one of these PBPs can not be obtained.

Although it has been shown that in *S. pneumoniae* at least three PBPs are encoded by mosaic genes in high-level resistant clinical isolates (PBP2x, PBP2b and PBP1a), the sole acquisition of a low-affinity PBP2x appears to be the prerequisite for high levels of resistance [61, 64]. In addition, whereas transformants of PBP2x and PBP2b constitute first-level resistance, second-level resistance is achieved upon transformation of PBP2a and/or PBP1b [65]. Nevertheless [61, 66], cell lysis does not occur as long as PBP2b is not inhibited. Therefore, strains containing an altered, low-affinity PBP2b are able to survive in the presence of  $\beta$ -lactams. It should be noted,



though, that PBP2b mutations are not selected with cefotaxime, since an important group of antibiotics, the expanded-spectrum cephalosporins, are [67] known not to interact with PBP2b. This makes PBP2x an excellent enzyme for performing binding studies with these antibiotics *in vivo*.

Nevertheless, since it has been shown [66] that the drug concentration at which interaction with PBP2b occurs can be more than 20 times higher than the MIC, it has been suggested that this tolerance may be the important first step towards resistance development *in vivo*. Indeed [68], transformation experiments using chromosomal DNA have shown that resistance and tolerance can be conferred simultaneously upon introduction of both PBP2b and PBP2x of resistant strains in a single transformation event.

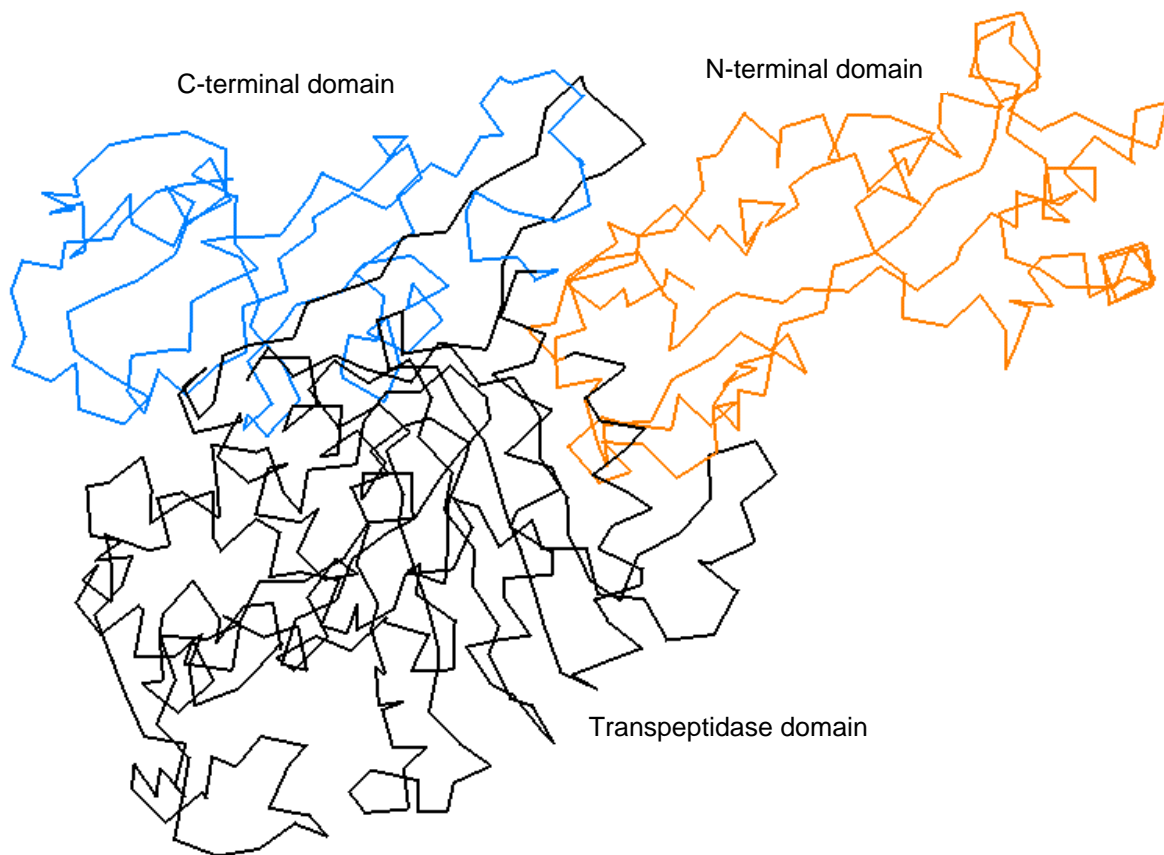
The situation has become even more complicated, since experiments with mutants resistant to piperacillin (a highly lytic  $\beta$ -lactam with high affinity to all pneumococcal PBPs) and mutants resistant to cefotaxime have led to the conclusion that non-PBP genes contribute to resistance as well [69, 70, 71]. Indeed [69,70], it could be shown that in the cefotaxime-resistant mutants resistance was linked to mutations in the *ciaH* gene encoding for histidine protein kinase, and PBP2x was the first PBP affected. On the other hand, PBP2b was the first PBP mutated in the piperacillin-resistant mutants, whereas a putative glycosyltransferase CpoA was shown [71] to act as the primary resistance determinant.

### 1.6.2 Structure of *Streptococcus pneumoniae* PBP2x

As mentioned in section 1.3.1 PBP2x is a 750-amino-acid class-B PBP (82.35 kDa) consisting of three modules. This protein is the first high- $M_r$  PBP whose three-dimensional structure has been crystallographically solved. Like other high- $M_r$  PBPs, PBP2x is anchored into the cytoplasmic membrane by an unprocessed N-terminal region. However [72], a soluble form, PBP2x\*, has been obtained upon deletion of the region encoding for residues 19-48 in the *pbpX* gene, and subsequent expression in *Escherichia coli*. This soluble form was crystallised in 1993 [73] and its structure was solved by X-ray crystallography to 3.5 Å resolution in 1996 [74]. However, in 1998 only its C $^{\alpha}$ -atom coordinates were deposited in the Protein Data Bank (PDB) [75] under the code name 1PMD.pdb.

As shown in Figure 1.10, the transpeptidase domain is localised between a sugar-tongue-shaped N-terminal domain (spanning residues 76 to 265) and a C-terminal domain (spanning residues 644 to 750). The C-terminal domain is connected to the transpeptidase domain via a 28-residue-long flexible loop.

As mentioned in sections 1.5.1.3 the transpeptidase or penicillin-binding domain is highly homologous to that of other PBPs and Class A, C and D  $\beta$ -lactamases. Indeed, the active site is constructed of the same structural motifs mentioned in section 1.5.1.2. : the tetrad S\*337TMK with S\* representing the active site serine, and the triads S395SN and K547SG.



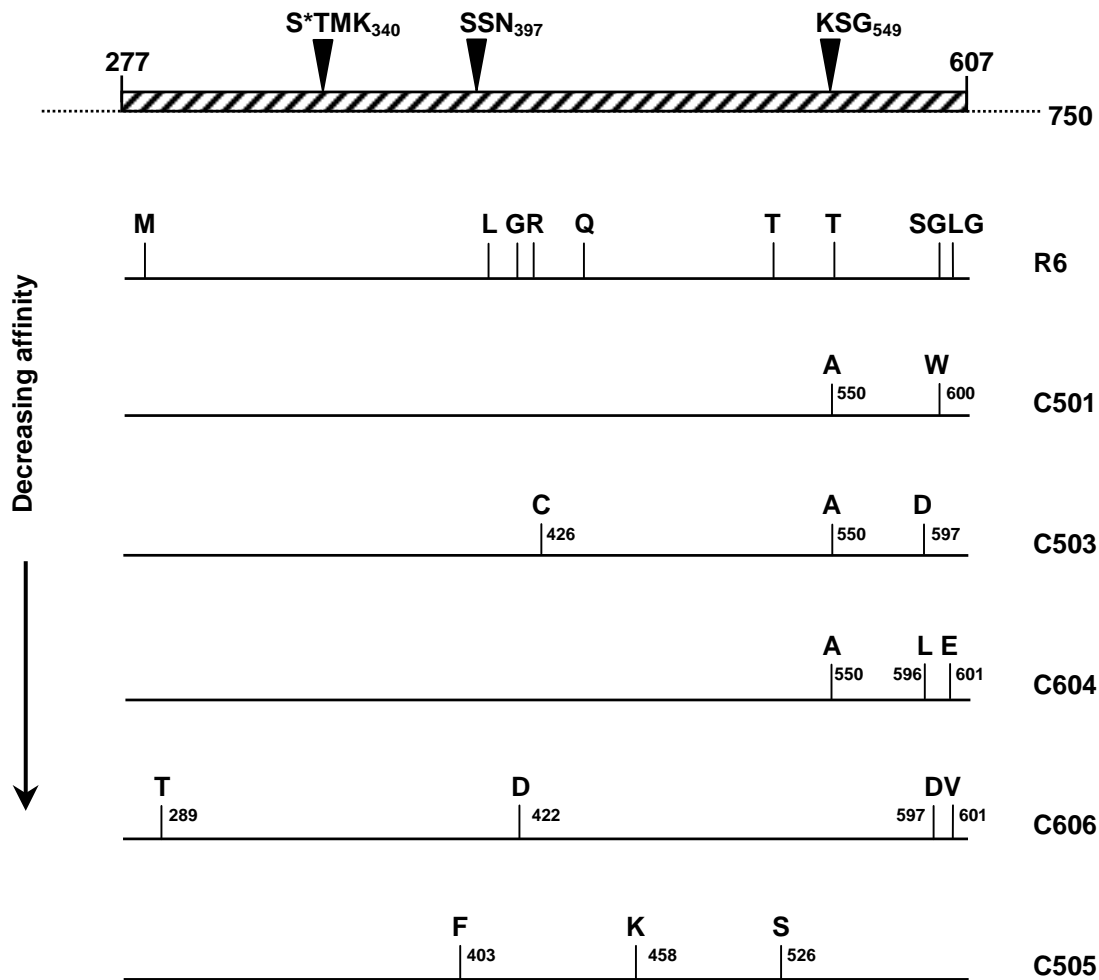
**Figure 1.10**  $C^\alpha$ -atoms of the 3.5 Å resolution structure of PBP2x\* (1PMD.pdb).

The function of the N-terminal and C-terminal domains is not known [28,74]. PBP2x does not appear to contain a transglycosylase domain [28, 62,74], though.

### 1.6.3 Amino acid mutations of PBP2x

Amino acid mutations in PBP2x are usually referenced to the penicillin-susceptible laboratory strain R6. Although naturally occurring mutants from clinically isolated drug resistant strains may have up to 100 substitutions in their amino acid sequence [76] with an average of 40 amino acid substitutions in the transpeptidase domain [77], substitutions T338A and Q552E are the most frequent [78]. Mutation of T338 into alanine has been associated [77] with the loss of a structurally bound water molecule near the active site forming hydrogen bonds with P335, T338, S571 and Y586. Infrared spectroscopy measurements have shown [79] that this water molecule has a stabilising effect on  $\beta$ -sheet B3 containing the K547SG motif and this structural determinant has been identified [77] as a primary resistance determinant in all Gram-positive bacteria.

However, other independent spontaneous amino acid substitutions [62,80] in five independent laboratory strains, have been shown to result in highly cefotaxime-resistant mutants as well (see Figure 1.11).



**Figure 1.11** Amino acid substitutions in PBP2x of five cefotaxime-resistant mutants [80].

The effects of mutations within or adjacent to the conserved amino acid motifs are confirmed by other data [81], although in some of the strains used here not only the *pbp2x* gene, but also the *pbp1a* and *pbp2b* genes appeared to be altered.

The single mutation of T550 (located in the 3D structure just below the K547TG motif) into alanine or glycine has been shown to result in a 6.4-fold increase of the MIC of cefotaxime [61]. In the same study, the T550G mutation has even been shown to result in a 12-fold increase of the MIC. Surprisingly, however, these mutations had practically no effect on the MICs for piperacillin and oxacillin.

Similarly, recent data [78] clearly show that the MIC for cefotaxime increases 18-fold upon mutation of T550 into alanine, whereas the same mutation seems to have practically no effect on the MIC of penicillin G.

The kinetic parameters of the mutations mentioned here have been summarised in Table 1.1 [82].

**Table 1.1** Comparison of kinetic parameters for the interactions between PBP2x-derivatives and  $\beta$ -lactams and thiolester substrate S2d.

PBP2x	benzylpenicillin $k_2/K$ ( $M^{-1}\cdot s^{-1}$ )	cefotaxime $k_2/K$ ( $M^{-1}\cdot s^{-1}$ )	S2d $k_{cat}/K_m$ ( $M^{-1}\cdot s^{-1}$ )
R6	58,000 $\pm$ 5,000	162,000 $\pm$ 4,000 <sup>1</sup>	2,600 $\pm$ 1,000
C503 <sub>G597D-T550A-R426C/ciaH/pbp2a</sub>	22,700 $\pm$ 300	1,700 $\pm$ 900	143 $\pm$ 27
C501 <sub>L600W-T550A/ciaH</sub>	10,000 $\pm$ 1,800	370 $\pm$ 16	130 $\pm$ 56
C505 <sub>T526S-L403F-Q458K/ciaH</sub>	52 $\pm$ 27	61 $\pm$ 25	< 50
C405 <sub>T526S-L403F/ciaH</sub>	52 $\pm$ 27	61 $\pm$ 25	< 50
R6 <sup>3</sup>	99,000 $\pm$ 12,000	209,000 $\pm$ 18,000	2,500 $\pm$ 200
T338A <sup>2</sup>	36,700	114,000	500
T550A <sup>3</sup>	112,000 $\pm$ 10,000	11,900 $\pm$ 900	1,360 $\pm$ 140

<sup>1</sup>data obtained in independent experiments

<sup>2</sup>from Mouz et al. [77]

<sup>3</sup>from Mouz et al. [78]

Altered substrate specificity is the biological price of mutating amino acids particularly in the active site region. It has been proposed [83-87] that this may have resulted in the heterogeneity of the cell wall composition.



## **2 Aim and scope of this thesis**

---





As explained in the Introduction, bacterial resistance to  $\beta$ -lactam antibiotics is a highly worrying problem in medical care. The decreasing affinity of PBPs to these compounds due to intrinsic resistance as described in sections 1.6 and 1.7 gives rise to an even more terrifying prospect.

*Streptococcus pneumoniae* is an important pathogenic agent in man. It is responsible for the majority of community-acquired pneumonia cases [88] and remains a non-negligible cause of mortality. Moreover, this bacterium is an important cause of septicaemia and otitis media, and one of the three most common pathogens in bacterial meningitis [89]. Although about 60 % of the entire human population is likely to carry *Streptococcus pneumoniae* in the nasopharynx [90], carriage is often accompanied by the development of antibodies against the *same* serotype. Infections, however, frequently occur upon the acquisition of a *different* serotype by inhalation of pneumococci from the upper respiratory tract.

*Streptococcus pneumoniae* R6 is one of the most penicillin-susceptible bacteria [40] (MIC < 10 ng/ml for benzylpenicillin). However, as explained in sections 1.5.2 and 1.6, this bacterial species is able to achieve resistance via gene-transfer resulting in the formation of mosaic genes. At present, many strains of *Streptococcus pneumoniae* are known worldwide [91], differing in their levels of resistance. Resistance spread has become a major problem due to the increased mobility of man. For example [92], it has been demonstrated that multiple clones of penicillin-resistant pneumococci have been introduced in the Netherlands, a country with a low prevalence of pneumococcal infection: in 1995 only 0.7 % of all pneumococci were intermediately resistant and another 0.4 % were highly resistant to penicillin. It has also been demonstrated that in 1997 in Germany, Austria and Switzerland about 95 % of the pneumococci were still penicillin-sensitive [93]. However, due to the increased mobility of man on the one hand, some clones spreading among the population in and outside hospitals on the other hand, it is urgent to gain insight into the effects of amino acid mutations of PBP2x resulting in resistance at the molecular level.

Since only the coordinates of the C $^{\alpha}$  atoms of a 3.5 Å resolution structure have been deposited in the PDB, however, in the first part of this dissertation an all-atom model of the transpeptidase domain will be generated, making use of homology modelling techniques.

In trying to explain experimental data for the naturally occurring T550A mutant, molecular dynamics simulations of the protein complexed with some penicillins and cephalosporins will be performed in the second part. Likewise, in order to shed light on the mechanistic consequences of the T526S point mutation present in the naturally occurring resistant C505 and C405 strains on the formation of Michaelis complexes molecular dynamics simulations will be performed of this mutant protein complexed with benzylpenicillin and cefotaxime.

## **3 Methodology**

---



## 3.1 Describing molecular systems

The central aim of Computer-Aided Drug Design is to obtain a quantitative understanding of drug-receptor interactions at the molecular level. For this purpose it is necessary to obtain detailed information about the conformations of both the ligand and the receptor in their bound states. Although the theory for describing a molecule's conformation and properties was already developed in the beginning of the nineteenth century, application to molecules consisting of more than two atoms has only become possible thanks to the revolution in computer-power during the past few decennia. Nevertheless, various methods differing in degree of approximation have been developed, since it is still not possible to give a quantum mechanical description of whole proteins.

### 3.1.1 Fundamentals

In order to be able to describe a molecular system, quantum mechanics makes use of wavefunctions, which contain *all there is to know* about such a system. One of the postulates of the theory is the statement that these functions  $\Psi$  exist. They are functions of the coordinates of all the particles and time:

$$\Psi(x_1, y_1, z_1, \dots, z_n, t) = \Psi(q, t) \quad (3.1)$$

$\Psi$  may be complex, and the probability that particle 1 can be found in a space  $V_1$  at time  $t$  is defined as  $\int \Psi^* \Psi d\tau$ , where

$$\int \Psi^* \Psi d\tau = 1 \quad (3.2)$$

The second postulate states that an operator exists for each physical observable and according to the third postulate the wavefunctions  $\Psi(q, t)$  are the solutions of the *time-*

*dependent* Schrödinger equation (relativistic effects on the electron mass are neglected):

$$\hat{H}(q,t) \Psi(q,t) = -\frac{\hbar}{i} \frac{d}{dt} \Psi(q,t) \quad (3.3)$$

where  $\hbar$  is the Dirac constant (or Planck's constant  $h$  divided by  $2\pi$ ) and  $\hat{H}$  is the so-called Hamilton operator (or Hamiltonian), containing the potential and kinetic energy of the system.

Whenever time is not explicitly contained in the Hamiltonian, equation (3.3) can be separated and solutions of the following form exist:

$$\Psi(q,t) = \Psi(q)\Phi(t) \quad (3.4)$$

Because the probability  $\Psi^*\Psi$  is independent of  $t$ , solutions given by (3.4) are called stationary states. The functions  $\Psi$  and  $\Phi$  are now solutions of

$$-\frac{\hbar}{i} \frac{d}{dt} \Phi(t) = E \Phi(t) \quad (3.5)$$

and

$$\hat{H}(q) \Psi(q) = E \Psi(q) \quad (3.6)$$

where  $E$  is the energy of the system for state  $\Psi(q,t)$ . Equation (3.6) is called the *time-independent* Schrödinger equation. For molecules, (3.6) can be written as follows:

$$\hat{H} \Psi_{e,n} = E_{\text{mol}} \Psi_{e,n} \quad (3.7)$$

where  $e$  and  $n$  are the electrons and the nuclei, respectively. Since equation (3.7) can only be solved for the hydrogen atom (and hydrogen-like ions), it is desirable to

separate the motions of the nuclei and the electrons. This is achieved by the Born-Oppenheimer approximation [94].

Approximated solutions of (3.7) can then be given by application of either the perturbation theory or the variation method. According to the latter method the energy of the system is calculated by solving the secular determinant:

$$|H_{rs} - W_i S_{rs}| = 0 \quad (3.8)$$

in which the variational integral  $H_{rs} = \int \Phi_r | \hat{H} | \Phi_s d\tau$  ( $\Phi_r$  and  $\Phi_s$  are trial functions for orbitals  $r$  and  $s$ ), the overlap integral  $S_{rs} = \int \Phi_r \Phi_s d\tau$ , and  $W_i$  is the eigenvalue or energy of state  $i$ . It can be shown, that the value of the variational integral  $W$  is always greater than the lowest possible energy of the system (variation theorem):

$$\frac{\int \Phi^* \hat{H} \Phi d\tau}{\int \Phi^* \Phi} \geq E_1 \quad (3.9)$$

So, when an expression for the wavefunction  $\Psi$  is given, information about the molecular system can be extracted with the Hamilton operator acting on it. There are four main approaches to calculating molecular properties. In order of decreasing accuracy these are: (1) *ab initio* methods, (2) semiempirical methods, (3) the density-functional method, and (4) the molecular mechanics method. In the following subsections (1) and (2) will be briefly described. Molecular mechanics methods will be described in section 3.2.

### 3.1.2 Quantum chemical methods

An *ab initio* (Latin for "from the beginning") calculation uses the correct molecular Hamiltonian and does not use any experimental data other than the values of the fundamental physical constants. In practice, this means solving the secular determinant (3.8), which becomes impossible for atoms containing more than two electrons.

### 3.1.2.1 *Ab initio* SCF MO methods

However, a key development in quantum chemistry has been the computation of accurate self-consistent-field (SCF) wavefunctions by Douglas Hartree and Vladimir Fock. With the Hartree-Fock method improved orbitals for each electron can be calculated in the *average field* generated by the other electrons:

$$\hat{F}\Phi_i = \varepsilon_i\Phi_i \quad (3.10)$$

where  $\hat{F}$  is a one-electron operator acting on the spinorbital for electron  $i$ , and the eigenvalue  $\varepsilon_i$  its corresponding energy value. A Hartree-Fock SCF calculation seeks the antisymmetrised product  $\Phi$  of one-electron functions that gives the minimum for the energy expression  $\int \Phi^* \hat{H} \Phi d\tau$ , where  $\hat{H}$  is the true Hamiltonian. Since properties like stability and reactivity of a compound can be explained in terms of s-, p-, or d-orbitals (valence theory) for which the functions are known, Roothaan [95] proposed in 1951 representing the Hartree-Fock orbitals as linear combinations of these basis functions, called atomic orbitals or AOs (LCAO-MO formalism). These AOs can be described by Slater-type orbitals (STOs) of the form  $X = \sum_i a_i e^{-c_i r}$  or Gaussian-type orbitals (GTOs) of the form  $X = r^{c_1} e^{-c_2 r}$ . Although STOs are more accurate in describing the electron cloud around an atomic nucleus, HF-SCF calculations on molecules with three or more atoms, evaluation of the many two-electron integrals (ab|cd) is impractical ("two electron integral problem"). For this reason, GTOs are generally used [96], in which the product of two Gaussians at different centers is equivalent to a single Gaussian function centred at a point between the two centers. Equation (3.10) now results in a set of Roothaan equations [97]:

$$\mathbf{F}\mathbf{c} = \mathbf{S}\mathbf{c}\varepsilon \quad (3.11)$$

where  $\mathbf{F}$  is the Fock matrix,  $\mathbf{S}$  is the overlap matrix,  $\mathbf{c}$  is an  $M \times M$  matrix composed of the orbital coefficients, and  $\varepsilon$  is an  $M \times M$  matrix of the orbital energies  $\varepsilon_i$ . The



Roothaan equations (3.11) have a non-trivial solution only if the following secular determinant is satisfied:

$$|\mathbf{F} - \epsilon_i \mathbf{S}| = 0 \quad (3.12)$$

The accuracy of the energy calculated depends strongly on the number of Gaussian functions describing the AOs (one-electron functions), or basis set. Because it is often insufficient to use only one STO or GTO for describing 1s, 2s, 2p<sub>x</sub>..... AOs, more functions are generally used. However, the more basis functions are used, the greater the number of Slater determinants that need to be solved, and the more time-costly the calculation. Therefore, it is more efficient to first optimise the structure using a simple basis set, before embarking on a more sophisticated calculation. The simplest basis set is STO-3G, in which each STO is replaced by three GTOs. For more sophisticated basis sets, it is useful to divide the orbitals into inner and outer orbitals. An example of such a split-valence basis set is the commonly used 6-31G basis set, which describes the inner orbitals by six Gaussian functions and the outer orbitals by two sets of basis functions, consisting of three functions and one function respectively.

More realistic, however, is to allow for deformation of orbitals by polarisation. Such a basis set is characterised by stars: one star if only the d-orbitals are allowed to deform (e.g. 4-31G\*), and two stars if the p-orbitals are allowed to deform as well (e.g. 6-31G\*\*).

Due to four reasons, the energy calculated by Hartree-Fock methods is always higher than the actual energy (see also equation (3.9)): (1) neglect of or incomplete treatment of electron correlation, (2) incompleteness of the basis set as described above, (3) neglect of relativistic effects, and (4) deviations from the Born-Oppenheimer approximation. In calculations on molecules without heavy atoms, however, (1) and (2) are the main sources of error. Therefore, an even better description of the equilibrium geometry is given by post-Hartree-Fock methods, like the Configuration-Interaction (CI), allowing for electron-electron interaction, and the Møller-Plesset [98] perturbation methods. Thus, it can be concluded that the term *ab initio* does not mean "100 % correct", since an *ab initio* SCF MO calculation uses the

approximation of taking  $\psi$  as an antisymmetrised product of one-electron spin-orbitals and uses a finite (and hence incomplete) basis set.

### 3.1.2.2 Semiempirical SCF MO methods

As pointed out in the previous subsection, *ab initio* calculation of polyatomic molecules can be a tedious task. In order to achieve reasonable approximations to *ab initio* results, however, the semiempirical methods have been developed, which only calculate the probability densities of the valence electrons. These methods fall into two categories: those using a Hamiltonian that is the sum of one-electron terms (e.g. the Hückel Method, which neglects differential overlap, the Extended Hückel Method [99], which does not neglect differential overlap, and the Pariser-Parr-Pople (PPP) method), and those using a Hamiltonian that includes both two-electron repulsion terms and one-electron terms.

Two-electron generalisations of the PPP method are the CNDO (complete neglect of differential overlap) [100], INDO (intermediate neglect of differential overlap) [101] and NDDO (neglect of diatomic differential overlap) [101] methods. These methods were developed with the aim of reproducing the *ab initio* SCF wave functions and properties as closely as possible. Whereas both the CNDO and INDO methods only include Coulomb-integrals describing the average repulsion between two electrons  $a$  and  $b$  in different valence spinorbitals  $\phi_i$  and  $\phi_j$  located on the *same* atom, NDDO also includes all two-center two electron interaction integrals describing the repulsion between two electrons located on *different* atoms. So, whereas both CNDO and INDO need only one parameter for each pair of atoms, NDDO makes use of 22 different two-center two-electron integrals [102]. Although this approach is the most correct, calculation of all integrals is very time-consuming. As a result, the benefit of incorporating electron correlation effects by adapting the integrals to experiment would vanish.

The general semiempirical methods developed by Dewar and coworkers, however, include various empirical parameters, thus speeding up calculation times considerably. Examples of these methods include MINDO (modified intermediate neglect of differential overlap) [103], MNDO (modified neglect of differential overlap) [104], AM1 (Austin Model 1) [105] and PM3 (parametric method 3) [106] are aimed at

having a theory that would give molecular binding energies with chemical accuracy (within 1 kcal/mol) and that could be used for large molecules without a prohibitive amount of calculation. These Dewar-type theories are parameterised so as to yield good values of the 25° C gas-phase standard enthalpy of formation  $\Delta H_{f,298}^0$ . For a molecule M consisting of atoms  $A_i$  this parameter is calculated from the molecular valence electronic energy  $U_{\text{val}}$  as follows:

$$\Delta H_{f,298,M(\text{g})}^0 = N_A U_{\text{val},e,M} - N_A \sum_i E_{\text{val},A_i} + \sum_i \Delta H_{f,298,A_i(\text{g})}^0 \quad (3.13)$$

where

$$U_{\text{val}} = E_{\text{el, val}} + V_{\text{cc}} \quad (3.14)$$

$$V_{\text{cc}} = \sum_{B>A} \sum_A [C_A C_B (s_A s_A | s_B s_B) + f_{AB}] \quad (3.15)$$

( $V_{\text{cc}}$  is the core-core repulsion energy and  $f_{AB}$  is an empirical function of the atomic distance  $R_{AB}$ ;  $C_A$  and  $C_B$  are the core charges of cores A and B). These methods differ essentially in the form of  $f_{AB}$ . Presently, AM1 is the most commonly used semiempirical method. It differs from MNDO by using sharper decreasing functions describing the Van der Waals repulsion [102] (see section 3.1.3). It has been parameterised for H, B, Al, C, Si, Ge, Sn, N, P, O, S, F, Cl, Br, I, Zn and Hg.

### 3.1.3 Force field methods

#### 3.1.3.1 Backgrounds

Although thanks to the explosion in computational power quantum chemical calculations have become more and more important during the past few decennia, today's computers are still not fast enough to give a complete quantum mechanical description of large molecules like proteins. Therefore, in order to be able to study the binding of ligands to large molecules like proteins, it is still necessary to make use of

methods making even greater approximations than the semiempirical quantum mechanical methods. Whereas the latter use a set of parameters describing only the core orbitals, with force field methods molecules are solely described by parameterised potential functions. These methods are not quantum mechanical methods, since force field methods do not deal with an electronic Hamiltonian or wave function or an electron density. Instead, molecules are treated as a set of masses (atoms) held together by springs (bonds). The interaction between these atoms is described with potential energy functions, which are constructed as a function of atomic coordinates, making use of parameters for describing bond-stretching and bond-bending and allowing for interactions between non-bonded atoms. So, whereas the quantum mechanical approach requires mapping out the potential energy surface for the molecule in question, the potential energy surface itself (in simplified form) is the starting point for the MM approach.

Generally, force-field methods describe the potential energy surface of molecules as a sum of bond stretching ( $V_{\text{str}}$ ), bond-angle bending ( $V_{\text{bend}}$ ), out-of-plane bending ( $V_{\text{oop}}$ ), internal rotation (torsion) about bonds ( $V_{\text{tors}}$ ), interactions between these kinds of motion ( $V_{\text{cross}}$ ), van der Waals attractions and repulsions between non-bonded atoms ( $V_{\text{vdW}}$ ), and electrostatic interactions between non-bonded atoms ( $V_{\text{el}}$ ):

$$V = V_{\text{str}} + V_{\text{bend}} + V_{\text{oop}} + V_{\text{tors}} + V_{\text{cross}} + V_{\text{vdW}} + V_{\text{el}} \quad (3.16)$$

In (3.16)  $V$  is only a measure of intramolecular strain relative to a hypothetical situation. By itself,  $V$  has no physical meaning. It is only the sum of components that matters. If there are other mechanisms affecting the energy, such as electrostatic repulsions or hydrogen bonding, these too may be added to the force field. Because each of the terms contained in (3.16) are described by analytical formulas (which will be described in section 3.1.3.3), MM methods are sometimes called *empirical-force-field* methods. In contrast to the quantum mechanical electronic single-point or geometry optimisation calculations, which need only input of atomic numbers of the atoms and their coordinates, in order to properly construct  $V$  force field methods also need specification of how the atoms are connected with each other. Although it is assumed, that force-field methods are inter transferable, artifacts are likely to be introduced when studying the behaviour of different molecules. This problem is

addressed in the spectroscopic force-field methods [107], in which the expression of the Born-Oppenheimer energy surface is expanded with a Taylor expansion.

Force-field methods can be divided into molecular mechanics (MM) methods and molecular dynamics (MD) methods. The former aim at optimising large molecules within the force field (see section 3.1.3.2), whereas the latter aim at studying the behaviour of molecules in time (see section 3.2.3).

### 3.1.3.2 Energy minimisation with force field methods

The aim of energy minimisation of a molecule is to locally explore the potential energy surface in order to find a point in the configuration space where the net force on each atom vanishes. These points are energetic minima. As mentioned in section 3.1.3.1 with molecular mechanics the starting conformation of a molecule is often chosen arbitrarily, which means that the potential energy of the molecule is not optimal. With energy minimisation methods the nearest minimum on the potential hypersurface (local minimum) can be found. This does not necessarily have to be the global minimum, though (in order to find this minimum molecular dynamics methods are usually used, which will be described in the next subsection).

The nearest local minimum can be found using iterative procedures, which calculate the gradient of the potential hypersurface, according to which new atomic positions are calculated of a lower-energy structure. Most methods apply the line search method, which seeks the nearest local minimum in the hyperpotential surface by stepwise generation of vectors, which represent the direction in which the energy decreases fastest. Since the direction in which the minimisation is continued is perpendicular to the search line at this point, each line search is orthogonal to the previous one. Some examples of minimisation algorithms include the steepest descent, the conjugate gradient, the Newton-Raphson and the simplex methods. Because here only the first two methods have been applied, they will be briefly described in the following. Both are gradient methods.

The steepest descent method starts with defining the direction of the local downhill gradient. Obviously, this method may be rather inefficient on energy surfaces having

narrow valleys. The program Discover, however, uses a slightly modified procedure, in that the position of the trial point is continuously updated along the direction of the gradient. This way line searches can be truncated, leading to a considerable decrease in the number of function evaluations. Nevertheless, the method remains rather inefficient due to the orthogonality of each new step. Because this is only useful at the beginning of the minimisation procedure (oscillation results when the minimum is approached), this method is often used for an initial relaxation of poorly defined structures.

Conjugate gradients, however, calculate a new direction vector by incorporating the gradient at the endpoint of the line search, which is added to the gradient from the previous direction. As a result, the resulting gradient is orthogonal to the previous one, and the next direction is conjugate to all previous directions. Because convergence toward the minimum is achieved more efficiently than with steepest descent, conjugate gradient is the method of choice for large molecular systems.

### 3.1.3.3 The Consistent Valence Force Field

Because of the fact that it has been developed for proteins, the Consistent Valence Force Field (CVFF) has been used throughout the work presented here. Nevertheless, this force field not only contains terms for describing the hydrogen-bonding interaction between amides [108], but also terms for describing the interactions of some small molecules.

In the following the analytical form of the terms given in equation (3.16) will be briefly described. Although implemented in the CVFF, the cross potentials ( $V_{\text{cross}}$ ), which represent couplings between deformations of internal coordinates, are not used in the work presented here. Although these terms may be required to accurately describe vibrational frequencies, the computational effort increases dramatically when calculating large molecules like proteins. Moreover, their behaviour can become unstable when the starting structure is energetically highly unfavourable.

**Bond stretching**

Generally, in the CVFF the bond length is described by a harmonic potential, which is based on Hooke's law. Like in all spectroscopic force fields, the potential is calculated as the force needed to stretch a particular bond between atoms  $i$  and  $j$  with respect to its equilibrium length  $b_0$ :

$$V_{\text{str}} = \sum_{i>j} k_{B,ij} (b_{ij} - b_{ij}^0)^2 \quad (3.17)$$

( $k_B$  = force constant of the bond length B [kcal.mol<sup>-1</sup>.rad<sup>-2</sup>]).

Although the form of (3.17) is relatively simple, it has been shown, that it suits very well for the description of large molecules, as long as many different atom types with specific parameters are used. However, in order to achieve higher accuracy in the treatment of *small molecules*, it is often more advantageously to define less atom types, but using more computationally expensive potential functions. Since the CVFF has been originally developed for proteins, however, the bond stretching potential is calculated by a simple quadratic form given in (3.17) with the definition of many different parameters.

Bond lengths can also be calculated more correctly by including Morse [109] potentials<sup>1</sup>. Because of the considerable extra computational cost of calculating these potentials, however, these potentials have been neglected throughout the work presented here.

**Bond bending**

The potential energy  $V_{\text{bend}}$  of bond bending is taken as the sum of potential energies  $V_{\text{bend},ijk}$  for bending each bond angle between atoms  $i$ ,  $j$  and  $k$  contained in the molecule. Like the stretching potential, the bending potential has a harmonic form as well:

---

<sup>1</sup> Morse potential:  $V = \sum_b D_b [1 - e^{-\alpha(b-b_0)}]^2$

$D_b$  = bond dissociation energy [kcal.mol<sup>-1</sup>]  
 $\alpha$  = Morse anharmonicity parameter  
 $b, b_0$  = see equation (3.17)

$$V_{\text{bend}} = \sum k_{B,ijk} (\theta_{ijk} - \theta_{ijk}^0)^2 \quad (3.18)$$

where

$\theta_{ijk}$  = the actual value of the bond angle between atoms  $i$ ,  $j$  and  $k$

$\theta_{ijk}^0$  = the equilibrium value of the bond angle between atoms  $i$ ,  $j$  and  $k$

### **Torsional energy**

The term  $V_{\text{tors}}$  is taken as the sum of terms  $V_{\text{tors},ijkl}$  over all 1,4 atom pairs:

$$V_{\text{tors}} = \sum k_{\tau} [1 + \cos(n\varphi_{ijkl} - \varphi_{ijkl}^0)] \quad (3.19)$$

where

$k_{\tau}$  = torsional barrier [kcal.mol<sup>-1</sup>]

$n$  = number of minima over 360° (periodicity)

$\varphi_{ijkl}$  = actual torsional angle between atoms  $i$ ,  $j$ ,  $k$  and  $l$

$\varphi_{ijkl}^0$  = reference value of torsional angle between atoms  $i$ ,  $j$ ,  $k$  and  $l$

### **Out-of-plane energy**

To ensure planarity of an  $sp^2$  hybridised carbon atom an out-of-plane bending term is added to the expression of the total energy. In the CVFF this out-of-plane term has the following form:

$$V_{\text{oop}} = \sum k_{\text{oop}} [1 + \cos(n\chi_{ijkl} - \chi_{ijkl}^0)] \quad (3.20)$$

where

$k_{\text{oop}}$  = out-of-plane torsional parameter

Since the functional form of  $V_{\text{oop},ijk}$  is similar to that of  $V_{\text{tors},ijkl}$  the out-of-plane term is often referred to as improper torsion. Like the torsional energy term it is defined by four atoms, but now the four atoms have to be positioned in the same plane.



An example is given by the carbonyl group. Planarity of the three atoms bonded to the carbonyl carbon is ensured by adding a  $V_{\text{oop},ijk}$  term to this carbon atom.

### ***Van der Waals interactions***

Whereas the energy terms mentioned above only account for *neighbouring* atoms contained in one and the same molecule, additional terms describing nonbonded and 1,4, 1,5, 1,6 ...atom-pair interactions are also required in the force field to describe molecular interaction as accurately as possible. These essentially nonbonded interactions are usually assumed to be of the same form as the interactions between rare gas atoms [110], i.e., a long range  $R^{-6}$  attraction due to induced dipole-induced dipole interaction (London dispersion interaction) and a short-range  $R^{-12}$  (Pauli) repulsion resulting from the overlap of electron clouds. Both types of interaction are contained in the Van der Waals energy term (the 1,2 and 1,3 Van der Waals and electrostatic interactions are considered to be implicitly included in the bond-stretching and bond-bending parameters described above). The resulting Van der Waals energy term is therefore usually expressed as the Lennard-Jones 12-6 potential:

$$V_{\text{vdW}} = \sum_{i < j} \varepsilon_{IJ} \left[ \left( \frac{R_{IJ}^*}{R_{ij}} \right)^{12} - 2 \left( \frac{R_{IJ}^*}{R_{ij}} \right)^6 \right] = 4 \sum_{i < j} \varepsilon_{IJ} \left[ \left( \frac{\sigma_{IJ}}{R_{ij}} \right)^{12} - 2 \left( \frac{\sigma_{IJ}}{R_{ij}} \right)^6 \right] \quad (3.21)$$

where  $R_{IJ}$  is the distance between atoms  $i$  and  $j$ ,  $\varepsilon_{IJ}$  is the value of  $V_{\text{vdW},ij}$  at the minimum in the interaction curve,  $R_{IJ}^*$  is the equilibrium distance between atoms  $i$  and  $j$ , and  $\sigma_{IJ}$  is the distance at which  $V_{\text{vdW},ij}$  is zero.

It is worth reiterating at this point that correlated terms like dispersion energy are completely omitted by Hartree-Fock level calculations [111] described in section 3.1.2.1. Thus, it is evident that also the Hartree-Fock potential energy surface contains some severe distortions in several regions. Although dispersion is everywhere attractive, depending on the system under consideration the highly anisotropic electrostatic term can be either positive or negative.

### ***Electrostatic interactions***

The electrostatic term  $V_{\text{el}}$  is taken as the sum of Coulombic interactions between the partial atomic point charges of the system under study, except for the 1,2 and 1,3 pairs (reason given above):

$$V_{\text{el}} = \sum_{i < j} \frac{Q_i Q_j}{\epsilon_r R_{ij}} \quad (3.22)$$

where  $Q_i$ ,  $Q_j$  are the atomic charges of atoms  $i$  and  $j$  respectively,  $\epsilon_r$  is the dielectric constant of the medium, and  $R_{ij}$  is the distance between atoms  $i$  and  $j$ .

Although very fast, one of the major drawbacks of using force fields is that one has to derive new parameters for each bond type not previously included in the parameter list. In the CVFF this especially accounts for the calculation of small molecules, since they may contain a large variety of atom types, whereas the CVFF was originally developed for studying proteins.

#### **3.1.4 Charges**

Of all the forces describing molecular interaction, the electrostatic are the most wide ranging. Depending on the charge distribution in the molecule, a three-dimensional molecular electrostatic potential (MEP) is generated around it. In principle, these MEPs can be calculated directly from the information contained in the wave function. For large molecules, however, the electrostatic properties of a molecule are described by making use of topologically derived atom-centered point charges. These charges are mainly based on the electronegativity of the atoms, but their values also depend on the connectivity with other atoms. The main drawback of using these charges is the fact that they are independent of molecular conformation. Still, most commercially available molecular modelling packages make use of topological methods for describing the charge distribution in large molecules [112].

---

Empirical force fields, which are able to incorporate conformation dependent charges are currently under development [113].

The topological charges in the molecular modelling program package InsightII [114] used here have especially been derived for proteins.

## 3.2 Exploring the molecular configurational space

### 3.2.1 Macroscopic versus microscopic properties

Both QM and MM methodologies give a simplified representation of the real *microscopic* world. In order to connect results derived from these methods with physically measurable *macroscopic* properties [115], it is necessary to include the central concept of *statistical mechanics* when describing a molecular system. This concept states that the most probable configuration of a microscopic system can be described by the Boltzmann distribution factor [116]:

$$\frac{n_j^*}{N} = \frac{e^{-\beta\varepsilon_j}}{\sum_j e^{-\beta\varepsilon_j}} \quad (3.23)$$

where  $n_j^*$  is the number of particles in the most probable (equilibrium) state,  $N$  is the total number of particles,  $\beta$  is  $1/kT$ ,  $\varepsilon_j$  is the energy of configuration  $j$ , and  $j$  is the total number of configurations. So, it is necessary to explore the configurational space of the system.

However, as described in the previous subsection, with *static* molecular mechanics methods usually only the nearest local minimum with energy  $U$  can be found, which is related to the physically measurable free energy  $G$  by the enthalpy  $H$  as follows [117]:

$$\Delta H = \Delta U - p\Delta V \quad (3.24)$$

$$\Delta G = \Delta H - T\Delta S \quad (3.25)$$

( $p$  is the pressure,  $V$  the volume of the system)

Since the volume change can be neglected under normal circumstances, calculation of the energies  $U$  of different conformations enables calculation of  $\Delta G$ . The reliability of the model can then be tested by correlation of the calculated  $\Delta G$  with experimentally obtained values.

Two general approaches have been developed [115] for exploring the configurational space:

(1) random search methods:

- Metropolis or Monte Carlo simulation
- Distance geometry

(2) dynamic simulation methods

- molecular dynamics simulation
- stochastic dynamics simulation

Because of their relevance in this thesis, both the Monte Carlo and molecular dynamics simulation methods will be described in more detail.

### 3.2.2 Metropolis or Monte Carlo simulations

The aim of the Monte Carlo (MC) simulation procedure is to produce a (canonical) configuration ensemble. The significance or probability of each of these configurations is calculated by the Boltzmann distribution given in equation (3.23). The procedure starts with a particular configuration  $r_s$ . New configurations are then generated by randomly displacing one or more atoms:

$$r_{s+1} = r_s + \Delta r \quad (3.26)$$

The configurational space can be fully sampled by taking a large number of successive displacements  $\Delta r$ . The new configuration is accepted if it has a lower energy. Else, its probability is weighted by the Boltzmann distribution:

$$P_{\Delta}(E) = e^{\left(\frac{-\Delta E}{k_B T}\right)} \quad (3.27)$$

where  $k_B$  is the Boltzmann constant ( $1.38 \times 10^{-23} \text{ J K}^{-1}$ ). Thus, a Boltzmann ensemble is generated by taking only relevant configurations into account. Because of the high bond energy introduced after random displacement of one or more atoms, the MC procedure is only suitable for small molecules.

### 3.2.3 Molecular dynamics simulations

The configurational space can also be explored by adding kinetic energy to the system, thereby enabling the molecule to cross the potential wells separating local minima on the potential energy hypersurface. With molecular dynamics (MD) simulations molecules are allowed to move by integrating Newton's equation of motion over time  $t$  for all the atoms contained in the system:

$$\mathbf{F}_i(t) = m_i \mathbf{a}_i(t) \quad (3.28)$$

$$-\frac{\partial \mathbf{V}}{\partial \mathbf{r}_i} = m_i \frac{\partial^2 \mathbf{r}_i}{\partial t_i^2} \quad (3.29)$$

where  $\mathbf{F}_i(t)$  is the force on atom  $i$ ,  $m_i$  is the mass of atom  $i$ ,  $\mathbf{a}_i$  is the acceleration on atom  $i$ , and  $\mathbf{V}$  is the potential energy of the system.

Before starting MD simulation of the molecule, minimisation is required (see section 3.1.3.2) in order to remove possible strain, which may cause the structure to blow up. The integration time step is chosen to lie in the order of the C-H bond stretching frequency, whose period is on the order of  $10^{-14}$  second. Mostly 1 femtosecond is taken for the integration time step.

The simulation of the minimised structure is started by giving each atom an arbitrary velocity in an arbitrary direction. Equation (3.28) is then solved for each atom.

According to the Verlet leapfrog algorithm [118], new atomic velocities are calculated at time  $t$  plus half the time interval, after which new atomic positions and new accelerations can be evaluated:

$$\mathbf{v}_i(t + \frac{1}{2}\Delta t) = \mathbf{v}_i(t - \frac{1}{2}\Delta t) + \Delta t \mathbf{a}_i(t) \quad (3.30)$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \frac{1}{2}\Delta t) \quad (3.31)$$

$$\mathbf{a}_i(t + \Delta t) = \frac{\mathbf{f}_i(t + \Delta t)}{m_i} \quad (3.32)$$

(  $\mathbf{f}_i(t + \Delta t)$  is evaluated from  $-\partial V/\partial \mathbf{r}_i$  at the new atom positions)

### 3.3 Comparative protein modelling

#### 3.3.1 Introduction

It has long been known, that knowledge of how ligands interact with proteins gives us valuable insights in the molecular backgrounds of many diseases. Unfortunately, insight into the three-dimensional aspects of drug-receptor interaction is still hindered by the fact that up to now, a comprehensive and accurate theory of protein folding based on physical and chemical principles is still lacking. This situation has led to the development of direct and indirect modelling methods in rational drug design. The former requires knowledge of the three-dimensional structure of the receptor (or at least of the active site), whereas with the latter methods the three-dimensional structure of the biological target is unknown.

Presently, about 12,000 protein structures are known three-dimensionally. Their structures have been deposited in the Protein Data Bank (PDB) [75]. Although the number of known 3D structures of targets is rapidly rising as a result of the rapid developments in both crystallography and multidimensional NMR-spectroscopy, this number of experimentally known protein structures is still less than 1% of the more than 200,000 unique protein amino acid sequences currently known [119]. Moreover, if one considers that the human genome encodes about 100,000 proteins [120], it is clear that there is still a wealth of structural information unknown to us.

Nevertheless, there will always remain a number of receptors, whose structure can not be determined with either of these techniques. For example, since a large number of proteins are insoluble or even cross membranes, it is frequently not possible to sufficiently purify them for crystallisation. This is especially true for the G-protein coupled receptors, which contain seven membrane-spanning helical structures. In cases where the protein is attached to the membrane by a membrane anchor, however, one has to rely on genetic techniques, in order to have the RNA code for a soluble protein and subsequently solve its structure by NMR techniques or protein crystallography.



Fortunately, though, it has long been recognised that whenever a clear amino acid sequence relationship between two proteins exists, it is likely that their three-dimensional structure and functions are also similar [121], which in turn has led to the recognition of the existence of a dense network of kinship relations between proteins. Thus, the inference of characteristics of a particular protein regarding structure and (or) function can be used in the prediction of shape or functional properties of related proteins. Therefore, this so-called comparative modelling technique bridges the gap between the three-dimensionally known proteins and the structurally unknown proteins. In fact, it remains the only modelling method that can provide models with a root mean square (rms) error below 2 Å [122].

Thanks to the explosive increase in computer power during the past few decennia, it has become possible to handle large databases protein sequences and structures, which has led to the rapidly emerging field of bioinformatics. When other genome projects are finished as well, the sequences of even more than 500,000 proteins [122] will be known. Since it has been estimated that there are about 1,000 different protein fold families, one third of which was already structurally defined in 1997 [120, 123, 124], it is to be expected that at least one member of each family will soon be structurally known, enabling us to generate complete 3D structures of every known protein sequence. In addition, with the complete sequence of the human genome being available soon, comparative protein modelling allows us to generate 3D models for most of the proteins in the human genome.

All comparative modelling methods have four sequential steps in common [125, 126]:

- a) identification of structurally or functionally related proteins with known 3D structures
- b) aligning these 3D structures with the target sequence and simultaneous choice of templates
- c) generation of a model for the target sequence by incorporation of the template structures using the alignment
- d) relaxation and validation of the newly generated protein model

The quality of the resulting protein model strongly depends on the alignment itself on the one hand, and the way how the 3D model is calculated from it on the other hand. With regard to the last step, four groups of methods can be distinguished:

- a) rigid-body assembly methods
- b) segment matching methods
- c) satisfaction of spatial restraints, and
- d) predicting loops and side chains on a given backbone

In the next few sections, each of these methods will be addressed separately. All of these methods have their advantages and disadvantages, but depending on the problem to be addressed, combinations of these methods may also be applied.

### **3.3.2 Alignment methods**

#### **3.3.2.1 Background**

As pointed out in section 3.3.1, the quality of a protein model generated by comparative protein modelling rises and falls with the quality of the alignment with homologous proteins of known 3D structure. The underlying idea of performing such sequence alignments is that proteins showing identical or related function, at least in their active site regions must show similar 3D architecture as well. From an evolutionary point of view, this means that functionally related proteins must have originated from a common ancestor, since selective pressure at the level of protein structures is known [127] to result from the need to maintain function, which in turn requires maintenance of the specific three-dimensional structure. Indeed [128], proteins consisting of at least eighty residues with more than 25 % sequence identity are very likely to be structurally and functionally similar, although structure may be conserved with even lower sequence identity [129].

The idea of aligning protein sequences dates back to 1966, when Fitch [130], and Cantor and Jukes [131] used symmetric matrices of scores tabulating element

---

relationships. An algorithm searching for fixed-length segments of each sequence was used in order to find those segments showing the highest sum of scores. Their idea has led to the development of various alignment techniques, which all have the following three aspects in common:

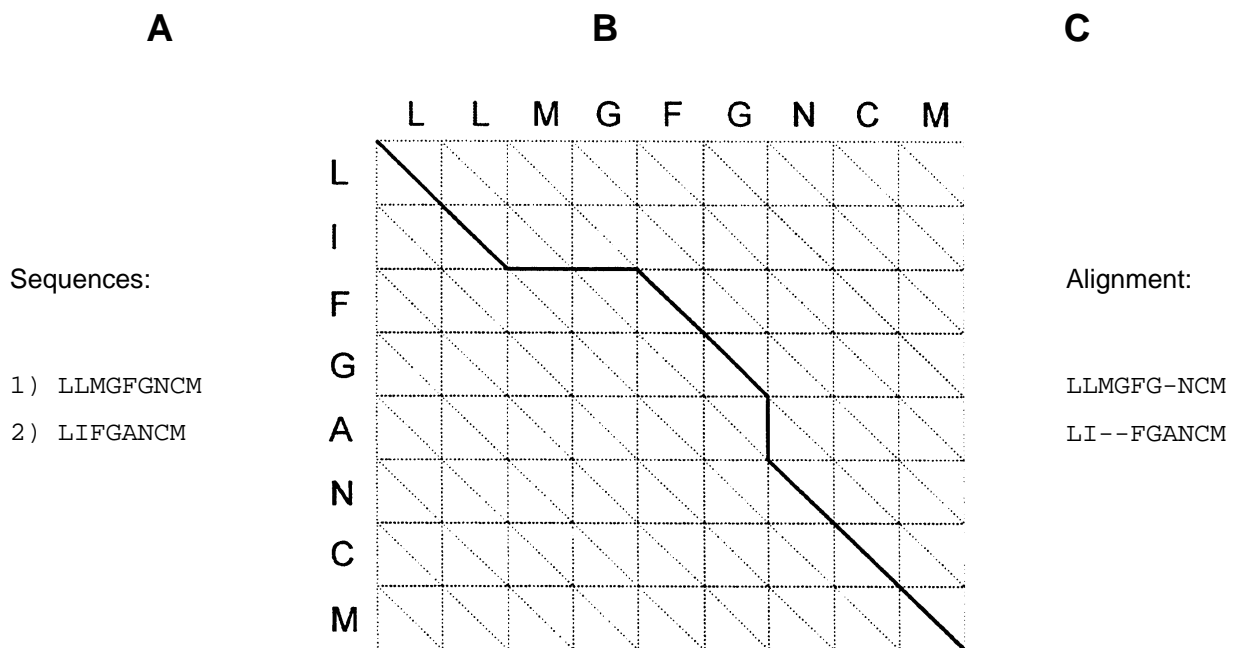
- (1) application of an algorithm for comparing sequences
  - Needleman-Wunsch algorithm (global alignments)
  - Smith-Waterman algorithm (local alignments)
- (2) application of a scoring function for amino acid mutations
- (3) application of a gap penalty function suppressing the insertion of gaps into the alignment

In the following, each of these aspects will be addressed separately.

### 3.3.2.2 Needleman-Wunsch algorithm for global alignments

The Needleman-Wunsch heuristic homology algorithm [132] is the first algorithm that was developed for describing *global* alignments of two proteins. The algorithm has been extended to produce an alignment for three proteins by Murata et al. [133]. The procedure starts with generating a comparison matrix of the sequences in question. Each matrix element is then assigned the value of 1 in the case of identical matches, and 0 for all nonidentical matches. The optimal path leading to the highest score is then found with dynamic programming. The method is illustrated in Figure 3.1.

The Needleman-Wunsch algorithm is the most widely used algorithm for performing global alignments. Its features can be extended when using substitution matrices, which will be described in subsection 3.3.2.4.



**Figure 3.1** Schematic representation of the pairwise global alignment procedure of Needleman and Wunsch [132], taking the sequences given in (A) as an example. Alignment of both sequences of varying length requires calculation of different ways through the matrix (B). By validating different kinds of amino acid combinations differently, the resulting alignment corresponds to the optimal route through the matrix.

### 3.3.2.3 Smith-Waterman algorithm for local alignments

When two sequences differ in length substantially, the Needleman-Wunsch algorithm gives rise to the insertion of numerous gaps (see section 3.3.2.5). Because possible similarities between longer subsequences may remain unidentified, this may eventually lead to very bad alignments. For example, it is not very useful to insert gaps into secondary structure elements.

A solution to this multi-dimensional problem can be provided by the Smith-Waterman algorithm [134] for *local* alignments. This algorithm finds the largest possible subsequences with high similarity score.

Both the Needleman-Wunsch and the Smith-Waterman algorithms are implemented in the two fast sequence alignment programs FASTA [135] and BLAST [136]. The aim of these programs is to find homologous proteins from a pool of structurally solved proteins.

### 3.3.2.4 Scoring functions for amino acid mutations

Since proteins are made up of 20 different amino acids, it is possible to derive matrices in which the elements represent a measure of equality between two amino acids. A first attempt to derive such a matrix was carried out by Dayhoff et al. [137] in 1978. They constructed phylogenetic trees from 71 groups of closely related proteins (> 85% pairwise sequence identity) and derived data for point accepted mutations (PAMs) per 100 residues. The matrix was updated in 1992 [138, 139], but the data had still been obtained from closely related sequences only.

Other matrices incorporating secondary structure information have been derived as well [e.g.140, 141, 142], another widely used set of matrices form the BLOSUM series [143]. These matrices have been constructed from the observation of amino acid substitutions in related proteins. Instead of only using proteins with higher than 85% identity, here far more distantly related proteins have been incorporated as well. Still [144], because it has been shown that the most important factors governing amino acid substitution during evolution, i.e. volume and hydrophobicity, the Dayhoff matrix remains the most widely used amino acid substitution matrix.

### 3.3.2.5 Gap penalty functions

When aligning two protein amino acid sequences, inevitably gaps are introduced. In order to minimise the number of gaps introduced into the alignment, gap penalty functions are used. An example of such a function is given below:

$$P = G_1L + G_2 \quad (3.33)$$

where

P = number of penalties

L = gap length

$G_1$  = parameter for extending an existing gap

$G_2$  = parameter for introducing a gap

Normally, the penalty for introducing a gap is 5-10 times higher than for extending an existing gap. In order to reduce the probability of introducing gaps in the alignment of secondary structure regions, it is possible to validate the penalty value in loop and secondary structure regions:

$$P_{ss} = Q(G_1L + G_2) \quad (3.34)$$

where Q is an arbitrary factor.

### 3.3.3 Rigid-body assembly approach

With rigid-body assembly methods a protein model is generated starting from a few core regions obtained by dissecting related structures [145, 146, 147]. Mostly the secondary structure regions are taken for these core regions, although some other structurally conserved regions may be included as well. These methods start with aligning the amino acid sequence of the model to be generated with related proteins of known 3D structure (see section 3.3.2). Subsequently, coordinates or the remaining loop regions can be generated by searching databases containing 3D protein structures for fragments, which (closely) match the given subsequence (threading methods). These fragments are then fitted between the two loose ends of the structure. Based on the fact that proteins have similar folds in their three-dimensional structure, Clothia et al. [148] have constructed the Structural Classification of Proteins (SCOP) database, which can be searched for protein segments matching a given amino acid sequence.

### 3.3.4 Segment matching methods

Another group of comparative modelling methods form the segment matching methods [149-153], which incorporate three-dimensional structure by searching databases for fragments that fit some common atoms. From these atoms (mostly backbone atoms or only the C<sup>α</sup>-atoms, see also section 3.3.6) the coordinates of

other atoms are then calculated. Levitt et al. [154] automated this approach for building complete proteins by systematically searching a database for fragments fitting a trace segment. For example, in order to generate a polypeptide backbone out of a given trace Claessens et al. [153] searched a database for substructures matching a given segment of it.

### 3.3.5 Satisfaction of spatial restraints

A third group of comparative modelling methods is modelling by satisfaction of spatial restraints [126, 155, 156] obtained from an alignment (distance geometry approach), which are based on the observation that similar 3D structures have similar interresidue distances. In these methods, the 3D shape is described with one or more matrices of intramolecular distances [157]. Alternatively, the parameter to be optimised may be a form of the Lennard-Jones potential function [158]. However, other constraints or restraints derived from many different sources may also be added to the homology-derived restraints. A review of the various methods satisfying spatial restraints is given in [125].

### 3.3.6 Protein modelling starting from C<sup>α</sup> coordinates

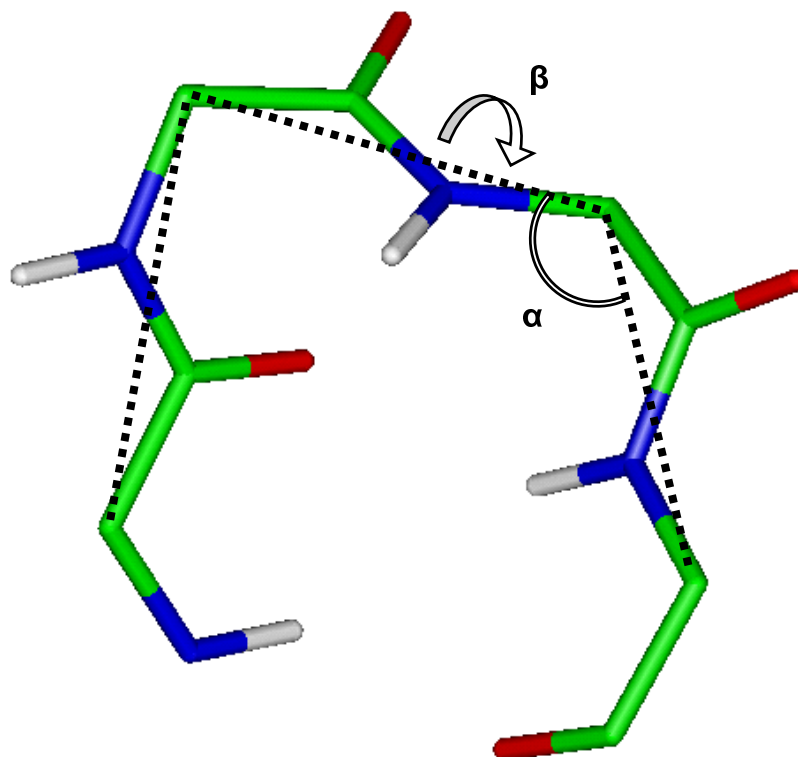
#### 3.3.6.1 Overview

Many methods have been developed in order to construct full protein coordinates from C<sup>α</sup>-atom information [149-151, 159-164]. Two groups of algorithms can be distinguished in these methods [165]: one based on the fitting of backbone fragments contained in a library (see also section 3.3.4) [149-151] and one based on geometric considerations [162-164]. Because they have been applied to some extent in the present work, some examples of the latter group, the virtual-bond chain method [166] and the dipole-path method [165] will be described in more detail.

### 3.3.6.2 The virtual-bond chain method

The virtual-bond chain method [166] makes use of the planarity of the peptide group. Since the distance between two constituent  $C^\alpha$ -atoms is therefore independent of  $\phi$  and  $\psi$ , this means that the conformational space of a polypeptide backbone can be described in terms of angles and dihedral angles between three and four subsequent  $C^\alpha$ -atoms, respectively.

By tabulating the values of  $\alpha$  and  $\beta$  (see Figure 3.1), the local structure of a closely related protein can be incorporated into a previously derived 3D model of a protein by folding the corresponding amino acid subchain.



**Figure 3.1** Schematic representation of a tripeptide showing the definition of the virtual bond angle  $\alpha$  and the virtual bond dihedral  $\beta$ .



### 3.3.6.3 The dipole-path method

Whereas the virtual-bond chain method describes the conformational space of the polypeptide chain by the angles between the from  $C^\alpha$ -atoms, the dipole-path method **[Error! Bookmark not defined.]** aims at reconstructing a backbone from  $C^\alpha$ -carbon coordinates based on peptide-group dipole alignment. The approach makes use of the assumption that, while a crude protein structure might be determined by hydrophobic packing forces, hydrogen-bonding interactions govern the arrangement of backbone atoms *within* a given virtual-bond chain conformation [167]. By representing the hydrogen-bonding network as electrostatic interactions between peptide-group dipoles, the problem of finding an optimal backbone conformation can be reduced to finding the best dipole alignment. Therefore, this method is based on an energetic rather than a purely geometric criterion.

## 3.4 Molecular docking methods

### 3.4.1 Backgrounds

Docking means determining the geometry of a receptor-ligand complex, prerequisites that the structure of the receptor is known [168]. This technique has become a key component of lead generation in structure-based design [169], since it has been possible to rapidly and accurately dock large numbers of small molecules into the binding site of a target protein, such that the compounds are rank-ordered with respect to their goodness of fit. However, docking methods can also be used for finding good starting geometries for running molecular dynamics simulations of protein structures complexed with ligands.

With docking two major problems need to be addressed:

- a) finding a computationally viable free-energy evaluation model that searches the complex energy landscape in order to find a minimum value
- b) development of an algorithm that efficiently searches a complex energy landscape.

On the other hand, the computational demands should be kept at a reasonable level, which is the reason why up to now, no docking program allows for full flexibility of both the ligand and the receptor. Since it is well known that ligand binding may involve a wide range of structural changes in the receptor protein, though, this problem is the main drawback of the presently available automatic docking programs.

A very promising concept in an attempt to reduce the search space in docking algorithms may be the recent scaling of side chain flexibilities [170]. Utilisation of such an amino acid flexibility scale should allow for inclusion of the side chain flexibility of a limited number of residues in the binding pocket.

The problem of efficiently searching the conformational space has given rise to two general approaches to the process of docking: the rigid-body approach and the flexible docking approach.

### 3.4.2 Rigid-body approach

The rigid-body approach is the most rigorous approach to docking problems. In this approach it is assumed that both the geometries of the ligand and the receptor remain unchanged by complex formation [169, 171]. Execution time is relatively rapid with this approach, since an algorithm need only search the six-dimensional space of rotation and translation. Although it has been shown that docking accuracy strongly increases when for both the receptor and the ligand the same conformation as in the bound state are used, simple scoring functions often fail to distinguish between near-native structures and far from native structures. Only in the case of small molecules this approach has been proven to be useful [168]. Moreover [172], since it has been shown that use of the geometries of the receptor and the ligand in bound conformation significantly increases the accuracy of the results, these conformations should be known at least to some extent before the docking experiment is carried out.

### 3.4.3 Flexible-ligand docking

#### 3.4.3.1 Introduction

A much more physiological approach in automatic docking is the flexible-ligand approach. The following techniques are currently in use:

- (a) Monte Carlo or molecular dynamics docking of complete molecules
- (b) in-site combinatorial search
- (c) ligand build-up
- (d) site mapping and fragment assembly

Because flexible docking strategies require an unbiased sampling of the allowed conformational space of the ligand and a concurrent exploration of its six rigid-body

degrees of freedom in the anisotropic environment of the receptor, a priori knowledge of the ligand conformation is not required.

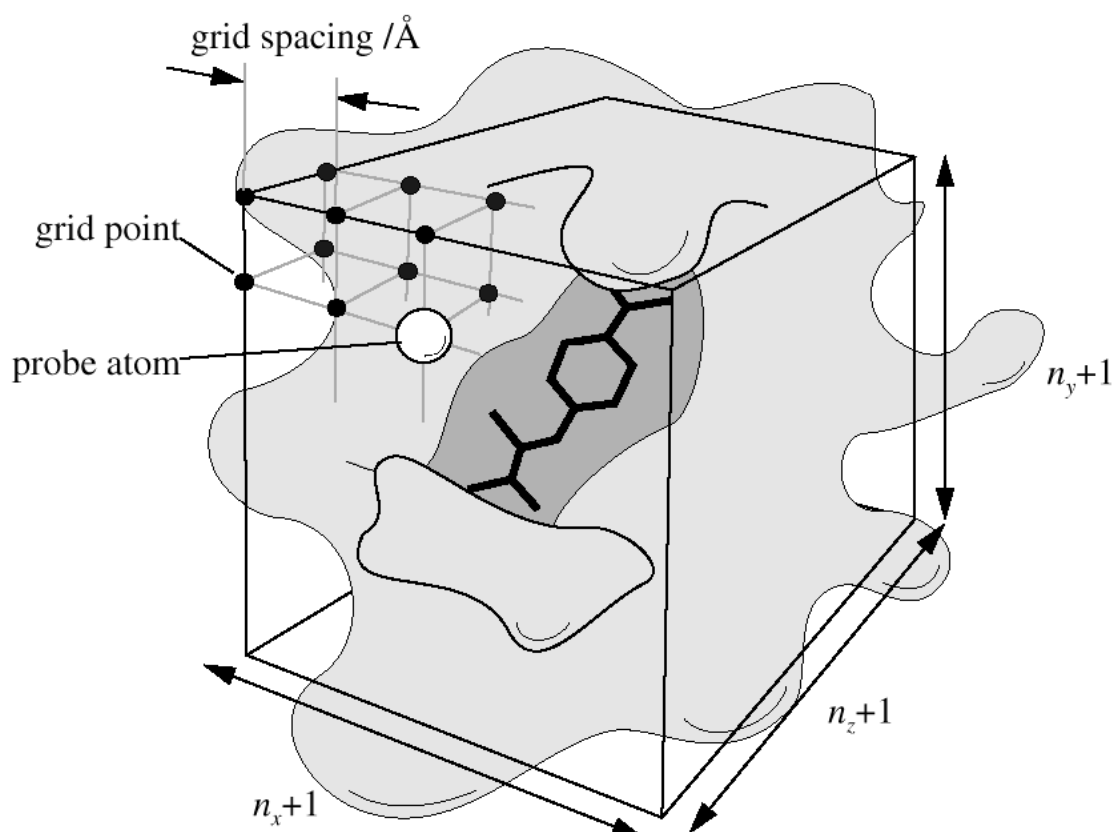
Four computational concepts play important roles in flexible docking:

- (a) shape descriptors consisting of a relatively small number of characteristic points on the surfaces of the two molecules in a complex
- (b) grid-based energy evaluation accounting for electrostatic and van der Waals interactions.
- (c) soft potentials that allow for some penetration of the surfaces being matched
- (d) local minimisation, although the results depend heavily on the starting conformation
- (e) multiple-copy techniques

#### **3.4.3.2 AutoDock 2.4**

AutoDock 2.4 [173] is a suite of the programs AutoDock, AutoGrid and AutoTors. These programs are designed to predict the bound conformation(s) of a flexible ligand to a macromolecular target of known structure, like an enzyme or DNA. They should be able to reproduce the x-ray crystallographically-determined positions of ligands having up to eight degrees of torsional freedom. In order to efficiently and quickly explore the configurational space this flexible docking program uses a Monte Carlo simulated annealing technique, combined with a rapid energy evaluation using precalculated grid based molecular affinity potentials. The force-field parameters are a subset of those used in AMBER [174].

First, the atom types present in the ligand are listed and a three-dimensional grid is calculated around the static protein (Figure 3.1).



**Figure 3.1** Schematic representation of the grid map placed around (the active site of) a protein in AutoDock [175].

Based on the work of Goodford [176], atomic affinity grid maps are calculated for each atom type present in the ligand being docked by storing the potential energy  $V$  of each probe atom at each grid point. Analogous to Equation 3.21 the van der Waals energies are calculated by a Lennard-Jones 6-12 potential:

$$E_{\text{vdw}}(r_{ij}) = \sum_{i < j, r_{ij} < r_{\text{cutoff}}} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \quad (3.35)$$

where  $r_{ij}$  is the distance between the interacting atoms  $i$  and  $j$ , and  $A_{ij}$  and  $B_{ij}$  are constants calculated from the well depth of the potential energy curve of these atoms. Hydrogen bonds are calculated with a 12-10 potential with different constants for  $A$

and B. Intermolecular hydrogen bonds can be calculated by adding polar hydrogen atoms.

In addition, the program AutoGrid calculates an electrostatic potential grid map, representing the Coulombic interactions between the macromolecule and a probe with elementary charge  $e$  (i.e. summation over all atoms of the macromolecule within a non-bonded cutoff of 8 Å radius). For this purpose partial atomic charges have to be assigned to the macromolecule, which are stored by adding an extra column to the usual PDB file. In order to model solvent screening a sigmoidal distance-dependent dielectric function according to Mehler and Solmajer [177] is used:

$$\epsilon(r) = A + \frac{B}{1 + ke^{-\lambda Br}} \quad (3.36)$$

where  $B = \epsilon_0 - A$ , with  $\epsilon_0 = 78.4$  (the dielectric constant of bulk water at 25 °C),  $A = -8.5525$ ,  $k = 7.7839$ , and  $\lambda = 0.003627 \text{ \AA}^{-1}$ . No distance cutoff is used for calculating electrostatic interactions.

After having prepared the molecular input files, a number of independent runs is calculated, each consisting of a sequence of constant temperature Monte Carlo simulated annealing cycles in order to explore the conformational space of the ligand. Quaternions (randomly placed axis in space around which rotation is performed) are used for rotation. Each cycle starts with a new temperature according to the following scheme:

$$T = gT_{i-1} \quad (3.37)$$

with  $g$  a factor between 0.85 and 0.95.

Thus, the ligand's current position, orientation and conformation are randomly changed. The resulting new state is then accepted or rejected upon comparison of the last move. According to equation 3.27, the probability  $P$  of acceptance of this new conformation depends on the energy difference  $\Delta E$  with its predecessor and the temperature  $T$ . The cycle ends if during a constant temperature cycle a user-specified number of accepted or rejected moves is reached. The next cycle starts with a new, randomly generated ligand conformation or with the minimum energy

---

state found during the previous cycle. In addition, it may be specified whether the next cycle starts at the last position of the previous cycle, or at a new randomly generated position.





## 4 Results

---



---

The work described in sections 4.1 to 4.3 is also described in [178].

## 4.1 Generation of all-atom coordinates

In section 1.6.2 it was already mentioned that penicillin-binding protein 2x of *Streptococcus pneumoniae* is the first high- $M_r$  PBP that has been crystallised [73], its structure solved to 3.5 Å resolution in 1996 [74], but that only its  $C^\alpha$  coordinates have been deposited in the PDB under the ID code 1PMD. Its structure is visualised in Figure 1.10 on page 24.

In order to perform MD studies of this protein complexed with various ligands, it was decided to use this poorly resolved structure to generate complete atom coordinates for the penicillin-sensitive transpeptidase domain and subsequently relax the structure with simulated annealing and molecular dynamics techniques.

Various methods for generating all-atom structures of proteins starting from  $C^\alpha$ -atom information have been described in section 3.3.6. In essence, these methods completely rely on the resolution of the  $C^\alpha$ -trace, which is not very useful if the crystal structure has been solved to a low resolution. Therefore, it is more desirable to incorporate some degree of optimisation of the structure during the process of generating full protein coordinates instead. For this purpose a new method was developed for generating full protein coordinates out of  $C^\alpha$ -atom information, combining local structural trace alignments and comparative side chain modelling with *ab initio* side chain modelling.

The program used for visualisation was InsightII Versions 97.0 and 98.0 [114]. Starting from 1PMD, an initial all-atom model of the transpeptidase domain of PBP2x was generated by dynamic programming upon simultaneously solving both the equations defining the bond angle constraint in terms of the peptide plane angles [179] and the virtual torsion constraint [166]. During the process of iteratively matching four consecutive  $C^\alpha$ -atoms with segments of known protein structures contained in a built-in database, frequently occurring conformations were statistically favoured. Peptide flips were introduced, because local secondary structure was not taken into account. Therefore, the trace atoms of the secondary structure elements of

---

the transpeptidase domain of PBP2x (defined as ranging from S266 to D616 as postulated by Pares et al. [74] were aligned locally structurally with the corresponding regions in five  $\beta$ -lactamases of known 3D structure (Figure 4.1). Similarly, some (parts of) loops could also be aligned. Using the Dayhoff PAM 250 substitution matrix [180] and the enhanced Needleman and Wunsch algorithm [132] for pairwise alignment as implemented in the Homology module of InsightII [114], the segments that were thus aligned are summarised in Table 4.1. A superposition of these parts with the crystal structure is shown in Figure 4.2.

```

BALI : DDFAKLEEQFDAQKLGIFALDTGT-NRTVAY-RP-----DERFAFAS
STAU : KELNDLEKKYNAHIGVYALDTKS-GKEVKF-NS-----DKRFAYAS
TEM1 : HPETLVKVKDAEDQLGARVGYIELDLSGKILESF-RP-----EERFPMMMS
BLT : PVSEKQLAEVVANTITPLMKAQSVFQGMVAVIY--QGKPHYITFGKAD-----IAANKPVTPQTLFELGS
CBL : AAKTEQIADIVNRTITPLMQEQAIQGMVAIIY--QGKPYFIFWGKAD-----IANNRPVTPQTLFELGS
PMD : DVYTTISPLQSFMETQMDAFQEKVKGKYMATLVSAKTGEILATIQRPFTDADTKEGITEDFVWRDILYQSNYEPGS
      260              280              300              320

BALI : TIKALTVGVLLQQKS--IEDLNQRITYTRDDLVNYPITEKHVDTGMTLKEADASLR-----
STAU : TSKAINSAILLEQVP--YNKLNKKVHINKDDIVAYSPILEKYVGKDITLKALIEASMT-----
TEM1 : TFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMTVRELCSAAIT-----
BLT : ISKTFTGVLGGDAIARGEISLDDAVTRYWPQLTGKQWQ-----GIRMLDLATYTAGGLPLQVPDEVTDNASLLRFY
CBL : VSKTFNGVLGGDAIARGEIKLSDPVTQYWPELTGKQWQ-----GISLLHLATYTAGGLPLQVPDDVTDKAALLRFY
PMD : TMKVMLAAIDN-----NTFP--GGEVFNSSELKIADATIRDWDVNEGL
      340              360              380

BALI : -----YSDNAAQNLIKQIG-----GPESLKKELRKIG-----DEVTNPERFEPELNEV
STAU : -----YSDNTANNKI IKEIG-----GIKKVKQRLKELG-----DKVTNPVRYEIELNYY
TEM1 : -----MSDNTAANLLLTIG-----GPKELTAFLNMG-----DHTVTRLDRWEPFELNEA
BLT : QNWQPQWK---PGTTRLIYANASIGLFGALAVKPSGMPYEQAMTTRVLKPIK-----LDHTWINVPKAEAAHY
CBL : QNWQPQWA---PGAKRLIYANSSIGLFGALAVKPSGMSYEEAMSKRVLHPIK-----LAHTWITVPQSEQKDY
PMD : TGGRMMTF---SQGFAHSSNVGMTLLEQKM-----GDATWLDYLNRFKFGVPTRFGLTDEYAGQLPADNIVNIA
      400              420              440

BALI : NPGETQDTST-ARALVTSLRAFALE-----DKLPSEKRELLIDWMKRN-TT---
STAU : SPKSKKDTST-PAAFGKTLNKLIAN-----GKLSKENKKFLLDLMLNN-KS---
TEM1 : IPNDERDTTM-PVAMATTLRKLITG-----ELLTLASRQQLIDWMEAD-KV---
BLT : AWGYRDGKAVRVSPGMLDAQAYGVKTNVQDMANWVMANMAPENV-----ADASLKQGIALLQSRYRWRI
CBL : AWGYREGKPVHVSPGQLDAEAYGVKSSSIDMTRWVQANMDASQV-----QEKTLQQGIELAQSRYRWRI
PMD : QSSFGQGISVTOQIMIRAFATAIANDGVMLEPKFISAIYDPNDQTARKSQKEIVGNPVSKDASLSTRNMVIVGTD---
      460              480              500              520

BALI : -----GDALIRAGVPDG-----WEVADKTGAAS-----YGTRN
STAU : -----GDTLIKDGVPKD-----YKVADKSGQAI-----TYASRN
TEM1 : -----AGPLLRSAIPAG-----WFIADKSGAGE-----R--GSRG
BLT : GSMYQGLGWEMLNWPVEANTVVEGSDSKVALAPLPAEAVNPPAPPVKASVWHKTGSTG-----GFGS
CBL : GDMYQGLGWEMLNWPLKADSIISGSDSKVALAALPAVEVNPPAPAVKASVWHKTGSTG-----GFGS
PMD : -----PVYGTMYNHSTGKPTVTVPGQNVALKSGTAQIADEKNGGYLVGLTDYIFSA
      540              560

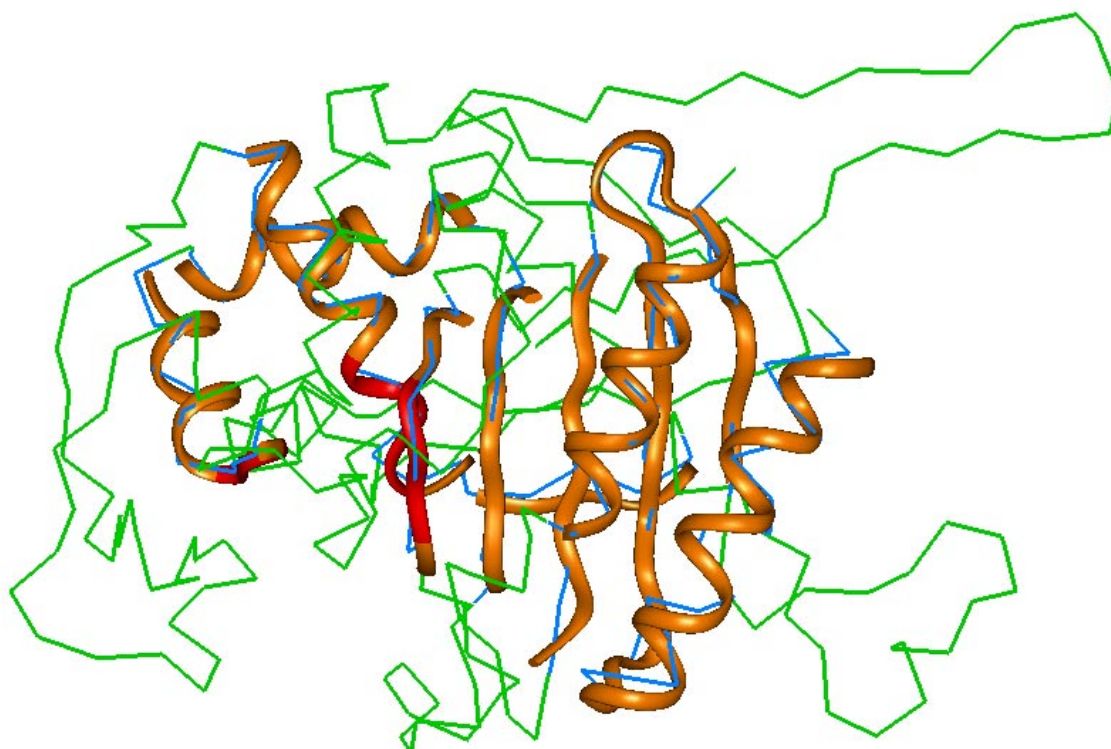
BALI : DIAIHWPPKGD-PVVLAVLSSRDKKDAKYDDKLI AEATKVMKALN
STAU : DVAFVYPKQSEPIVLVIFTNKDNKSDKPNDKL ISETAKSVMKEF
TEM1 : IIAALGPDGKP-SRIVVIYTTGSQA-TM---DERNRQIAEIGASLIKHW
BLT : YVAFIP---EK-QIGIVMLANTSY-----PNPARVEAAAYHILEAL
CBL : YVAFVP---EK-NLGI VMLANKSY-----PNPVRVEAAWRILEKLC
PMD : VSMSPA---ENPDFLLYVTVQDPEHYSGIQLG EFANPILERASAMKDSL N
      580              600

```

**Figure 4.1** Alignment of the transpeptidase domain of PBP2x<sup>\*</sup> with the five  $\beta$ -lactamases of known 3D structure.

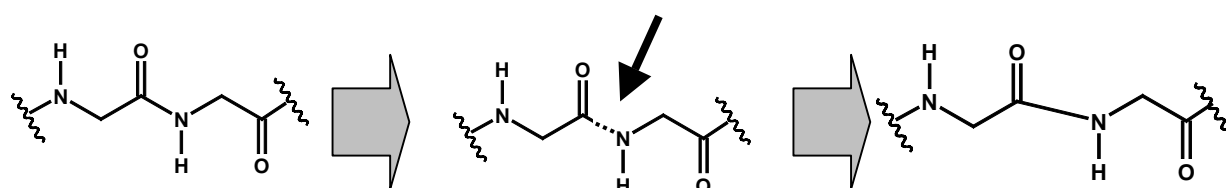
**Table 4.1** *Regions for which the C<sup>α</sup>-atoms were aligned with those of the β-lactamases, as well as their homology and RMSD values.*

Region	Length	β-Lactamase	Homology	RMSD
S <sub>267</sub> - K <sub>287</sub>	21	2BLT	-4.29	0.96
Y <sub>288</sub> - T <sub>305</sub>	18	1BTL	4.44	0.77
Q <sub>330</sub> - E <sub>334</sub>	5	1CBL	26.00	0.49
P <sub>335</sub> - N <sub>350</sub>	16	4BLM	19.38	0.54
H <sub>394</sub> - K <sub>406</sub>	12	4BLM	8.46	0.43
A <sub>410</sub> - K <sub>420</sub>	11	1BTL	7.27	0.51
T <sub>459</sub> - A <sub>469</sub>	11	1BTL	4.55	0.55
A <sub>508</sub> - L <sub>517</sub>	10	1BLT	6.00	0.52
V <sub>544</sub> - A <sub>551</sub>	8	4BLM	20.00	0.83
I <sub>569</sub> - S <sub>576</sub>	8	4BLM	3.75	0.68
F <sub>583</sub> - Q <sub>591</sub>	9	4BLM	3.33	0.82
E <sub>602</sub> - K <sub>615</sub>	14	1BTL	-3.57	0.60



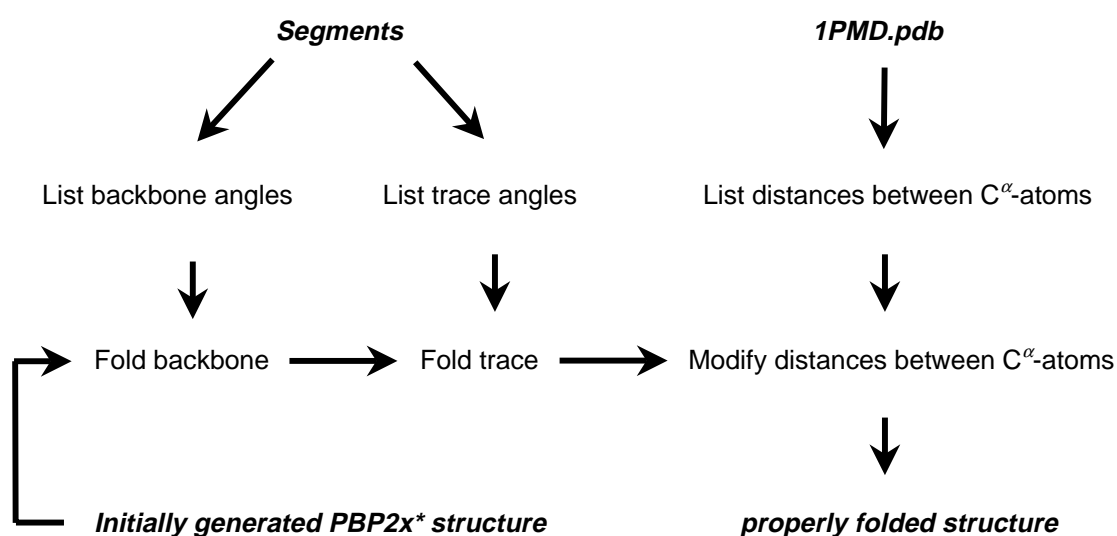
**Figure 4.2** *Superposition of 1PMD.pdb onto the total structure composed of fragments obtained by alignment of 1PMD.pdb with the 5 β-lactamases of known 3D structure (shown as ribbon). The corresponding regions in 1PMD.pdb are coloured blue. The parts of the ribbon covering the active site residues are coloured red.*

The geometry of the thus created segments was inserted into the structure generated by Biopolymer, by subsequently incorporating the angles and dihedrals between the main chain heavy atoms of the newly generated segments, refolding the trace according to the X-ray structure and finally modifying the distance between two consecutive  $C^\alpha$  atoms of the modified regions. The last step of the procedure was performed by subsequently breaking the peptide bonds connecting two amino acids, adjusting the distance between two consecutive  $C^\alpha$  atoms and finally recovering the peptide bond again (see Figure 4.3). Afterward, the trace geometry of the X-ray structure was completely recovered and the resulting structure was minimised.



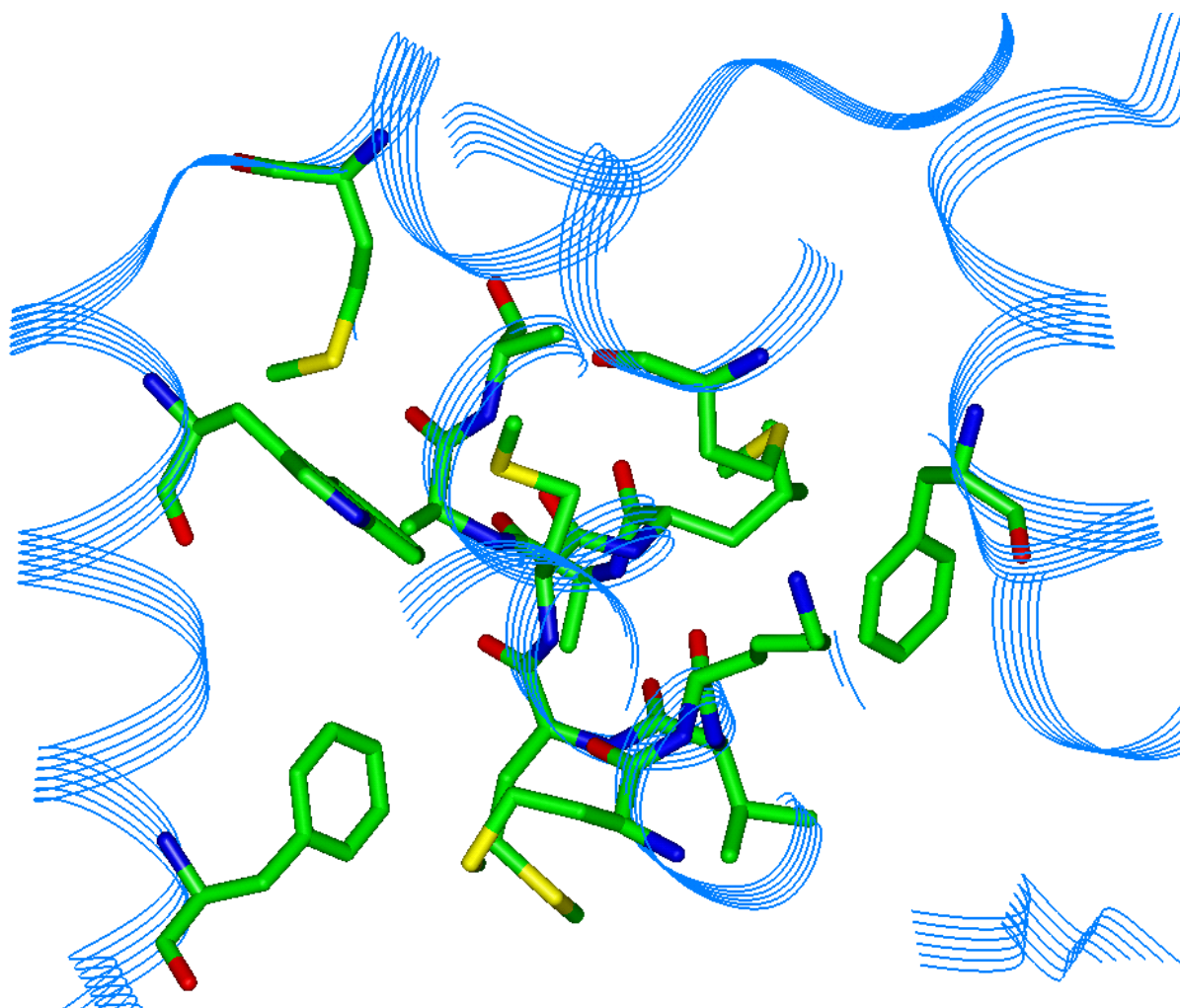
**Figure 4.3** Schematic representation of the readjustment of the distances between two subsequent  $C^\alpha$  atoms.

The whole procedure of (locally) modifying the backbone geometry of a previously generated protein structure according to the corresponding region of a template structure is shown in Figure 4.4.



**Figure 4.4** Flow diagram for the incorporation of the backbone geometry of the homology derived segments of 1PMD with  $\beta$ -lactamases of known three-dimensional structure.

In order to take account of as much structural information as available for the generation of an MD starting structure, the side chains of the methionine cluster next to the active site cluster (comprised of M339 to A347, F392, M400, M407, W412, and F419) were manually adjusted according to the geometry as displayed in the article by Pares et al. [74]. A view of the resulting conformation for this region is shown in Figure 4.5.



**Figure 4.5** *View of the resulting conformation of the methionine cluster next to the active site after manual adjustment of the dihedral angles of the side chains according to Figure 1 in [74].*



## 4.2 Relaxation of the transpeptidase domain

### 4.2.1 General comments

Using the program package Discover (version 97.0), the transpeptidase domain was relaxed with various molecular dynamics simulations. Since the positions and velocities calculated by equations 3.30 to 3.32 on page 52 are half a timestep out of synchrony, the Verlet velocity algorithm as implemented in Discover is used for calculating new positions, velocities and accelerations:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{\Delta t^2 \mathbf{a}_i(t)}{2} \quad (4.1)$$

$$\mathbf{a}_i(t + \Delta t) = \frac{\mathbf{f}_i(t + \Delta t)}{m_i} \quad (4.2)$$

$$\mathbf{v}_i(t + \frac{1}{2} \Delta t) = \mathbf{v}_i(t) + \frac{1}{2} \Delta t [\mathbf{a}_i(t) + \mathbf{a}_i(t + \Delta t)] \quad (4.3)$$

### 4.2.2 Relaxation strategy

For several reasons, application of a standard tethering force was found to be unsuitable during the side chain relaxation procedure. The underlying cause was the fact that, in contrast to standard relaxation of homology modelled protein structures, here it was apt to stick to the sole experimental information that was available, i.e. the C<sup>α</sup>-carbon coordinates. Since it was desirable to slowly relax the 3D homology derived regions in a positionally constrained framework made up by the remaining regions, while still allowing for enough flexibility, different force constants were needed. These criteria are met upon application of a *restraining* force. Equations 4.4a-b and 4.5 nicely explain how force constants are incorporated into the total

energy of the system in question when applying a tethering force and restraining force.

Tethering force:

$$\Delta E = K \cdot \text{RMSD} \quad (4.4a)$$

$$\Delta E = K \sqrt{\frac{\sum_{i=1}^N \Delta x_i^2 + \Delta y_i^2 + \Delta z_i^2}{N}} \quad (4.4b)$$

Restraining force:

$$\Delta E = \sum_{i=1}^N \left( K_i \cdot \sqrt{\Delta x_i^2 + \Delta y_i^2 + \Delta z_i^2} \right) \quad (4.5)$$

where

$E$  = internal energy of the system

$K$  = tethering force constant

$N$  = number of atoms to be tethered

$i$  = atom index

$x, y, z$  = cartesian atom coordinates

Whereas with tethering a *single* force constant is applied to all the atoms to be tethered and the resulting energy penalty is calculated by multiplication of the RMSD value, with restraining each atom can be restricted in its movement individually with different force constants.

Therefore, it was decided to run several MD simulations of the transpeptidase domain, during which the homology derived regions were slowly relaxed, while restraining the C<sup>α</sup>-carbon atoms in the remaining regions to their respective X-ray structure positions, and keeping both the C-terminal and the N-terminal domains fixed.

### 4.2.3 Relaxation of the backbone

However, due to the *local* alignment procedure of the secondary structure elements described above, the hydrogen bonding network of the five-stranded  $\beta$ -sheet was not optimal. Therefore, before embarking on MD simulations of the transpeptidase domain as a whole, the hydrogen bonding network resulting from the backbone N-H and C=O groups atoms had to be optimised. For this purpose all of the amino acids of PBP2x\* were transformed into alanine residues, except for the proline and glycine residues. This was performed by running the BCL script *Auto\_replace.bcl* given in Appendix F. Now, the only hydrogen bonds present in the structure were resulting from the N-H and C=O groups of the peptide moieties. Of the resulting poly-Ala-Gly-Pro structure a simulated annealing MD run at 1000 K was performed, while applying a *restraining* force of 10 - 15 kcal/mol/Å on the C $^{\alpha}$ -atoms. In addition, for all non-proline residues the omega angles were forced to *trans* conformation using 5 kcal/mole. After 20 ps initialisation and 80 ps simulation at this temperature, the system was cooled down to 0 K. Then the system was heated to 1000 K in 10 ps and cooled down again. It was frequently observed that the complete peptide group between two constituent C $^{\alpha}$ -atoms was rotated over 180°, while retaining the *trans* conformation. With the exception of P308, none of the peptide moieties was changed into *cis* conformation as was shown by PROCHECK [181]. With this procedure, 10 structures of the poly-Ala-Gly-Pro protein were obtained, and finally minimised (500 cycles steepest descent, followed by conjugate gradients). A superposition of these structures nicely showed that some regions in the protein are more flexible than other regions.

The lowest energy conformation was taken as starting conformation for performing the next simulated annealing calculation, in which the system was cooled from 600 K to 100 K. The restraining force constant was held constant, while the forcing constant applied to retain *trans* conformations (except for the proline residues) was reduced to 2 kcal/mol. The results obtained from these calculations are summarised in Table 4.2.

**Table 4. 2** *Analysis of the various simulated annealing MD simulations of the penicillin-binding domain of PBP2x\*.*

Temperature (K)	restrain / tether	force	simulation time (ps)	internal energy (kcal)	RMS trace (full trans- peptidase domain)	RMS trace (aligned regions)	%most favoured regions	%dis- allowed regions	main chain bonds	distorted main chain angles	planar groups
1000	restrain	10 - 15	390	2030.5	0.40	0.12	78.6	1.0	0	0	0
600	restrain	10 - 15	220	2018.1	0.41	0.17	78.0	1.0	0	0	0

#### 4.2.4 Relaxation of the full transpeptidase domain

In order to relax the whole transpeptidase domain, the native amino acid composition of the transpeptidase domain was fully recovered. By not recovering the natural amino acid compositions for the C-terminal and N-terminal domains, unrestricted movement of the side chains of the transpeptidase domain was assured during the MD simulations of the transpeptidase domain. After minimisation of the side chains a protein structure was obtained in which the backbone conformation originated from the previous run. After minimisation the resulting structure was taken as template for MD simulations of the transpeptidase domain.

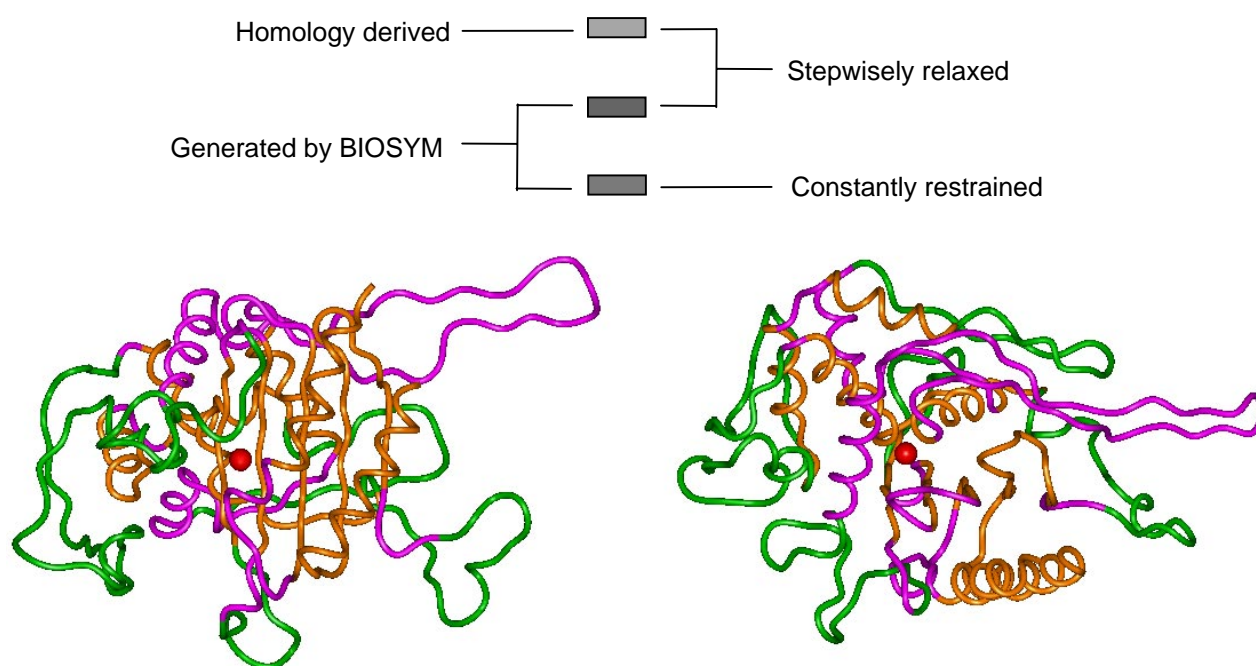
First, a simulated annealing calculation was performed at 610 K, during which the atoms mentioned before were *restrained* to their corresponding positions in the template using 2 - 5 kcal/mol. After 80 ps (including 20 ps initialisation) the system was cooled down to 310 K. Then it was allowed to relax during 15 ps, its conformation was saved, and subsequently it was again heated to 610 K during another 15 ps (the input file is given in Appendix J). By repeating this procedure several times, five structures were obtained, which were finally minimised. The best conformation was taken as a starting conformation for a series of MD simulations at 310 K of 220 ps each, with the aim of relaxing the transpeptidase domain as much as possible, hereby making use of the possibility of applying different force constants upon using a *restraining* force. At this point the presence of the structural water molecule [182] was included in the MD simulations. Indeed, a water molecule could be easily placed in the hole between P335, T338, Y586 and S571. During the following simulations the water molecule was kept in place by restraining the intermolecular hydrogen bonds with these four amino acids with 1.0 – 2.0 kcal/ mol/Å.

Furthermore, the transpeptidase domain was divided into two sets of regions. These are colour-coded in Figure 4.1. The green coloured regions represent those in which the C<sup>α</sup>-atoms were restrained to their positions in the template structure using 1.0 to 2.0 kcal/mole/Å in all simulations. These regions are listed in Table 4.3. The orange and magenta coloured regions represent those which were completely relaxed by performing a series of MD simulations, in which the force constant on the C<sup>α</sup>-atoms

was decreased from 1.0 - 2.0 kcal/mol/Å to 0 kcal/mol/Å in steps of 0.2 - 0.4 kcal/mol/Å.

**Table 4.3** *Regions in which the C<sup>α</sup>-atoms have been restrained with a constant restraining force of 1.0 - 2.0 kcal/mol/Å in all molecular dynamics simulations of the transpeptidase domain.*

F310 – Y333	F419 – D440
T352 – F392	V516 – S576



**Figure 4.1** *Ribbon representation of the transpeptidase domain of PBP2x\*. The structural water molecule is shown as a red dot. The regions in which the C<sup>α</sup>-atoms have been constantly restrained to their corresponding positions in the crystal structure are shown in green. Both the orange coloured and magenta coloured regions have been slowly relaxed by gradually decreasing the restraining force constant to zero. For further details see text.*

The underlying idea of dividing the transpeptidase domain into two sets of regions was to be able to completely relax the homology derived regions as much as possible, but within a framework of the other regions, of which only the C<sup>α</sup>-atoms

---

were experimentally known. In order to retain the accessibility of the active site, though, it was necessary to constrain the C<sup>α</sup>-atoms of B3 and B4 of the five-stranded β-sheet to their X-ray coordinates as well. During these simulations, the heavy atoms of the side chains contained in the methionine cluster were also relaxed stepwisely. Each of these six 220 ps MD simulations was started with 20 ps initialisation time. Afterward, every 20 ps the conformation was saved, all of which were subjected to a minimisation procedure afterwards (500 steps of steepest descent minimisation, followed by conjugate gradients minimisation until the energy was less than 1.00 kcal/mole). One of the ten resulting conformations was chosen as starting conformation for the next MD run (criteria used: low value of internal energy, low RMS deviation of the trace, high percentage of amino acids in most favoured regions, long enough a simulation time). In the last MD simulation the orange coloured regions were completely free to move within the framework of the large loops surrounding the transpeptidase domain. Since it was observed that the last helix of the transpeptidase domain (containing a proline residue) was unfolding when no restraining force was applied at all, in the last simulation the C<sup>α</sup>-atoms of this helix were restrained with a force constant of 0.2 – 0.4 kcal/mole/Å. A summary of the most important parameters obtained for the various MD simulations is given in Table 4.4.

**Table 4. 4** Analysis of the various MD simulations of the penicillin-binding domain of PBP2x\*.

Temperature (K)	restrain / tether	force	simulation time (ps)	internal energy (kcal)	RMS trace (full trans-peptidase domain)	RMS trace (aligned regions)	%most favoured regions	%dis-allowed regions	main chain bonds	distorted main chain angles	planar groups
610	restrain	2 – 5	220	2425.5	0.85	0.57	80.6	1.0	10	0	3
310	restrain	1.0 - 2.0	220	2395.0	0.93	0.74	79.6	1.0	6	0	3
310	restrain	0.8 - 1.6	220	2300.6	1.23	0.96	80.6	0.7	9	0	2
310	restrain	0.6 - 1.2	220	2292.6	1.32	1.02	82.2	0.7	8	0	2
310	restrain	0.4 - 0.8	220	2282.5	1.45	1.05	82.2	0.7	8	0	0
310	restrain	0.2 - 0.4	220	2288.0	1.52	1.00	82.2	0.7	7	1	2
310	restrain	0	220	2260.1	1.64	1.29	79.3	1.0	8	2	4



---

Since in the meantime the crystal structure has been solved to a resolution of 2.4 Å [183], making localisation of the side chains possible, at this point, the model was sent to Dr Dideberg for verification. Almost all of the side chains appeared to be in the correct orientation [184], except for F359 and M386. Major differences appeared only in the loops.

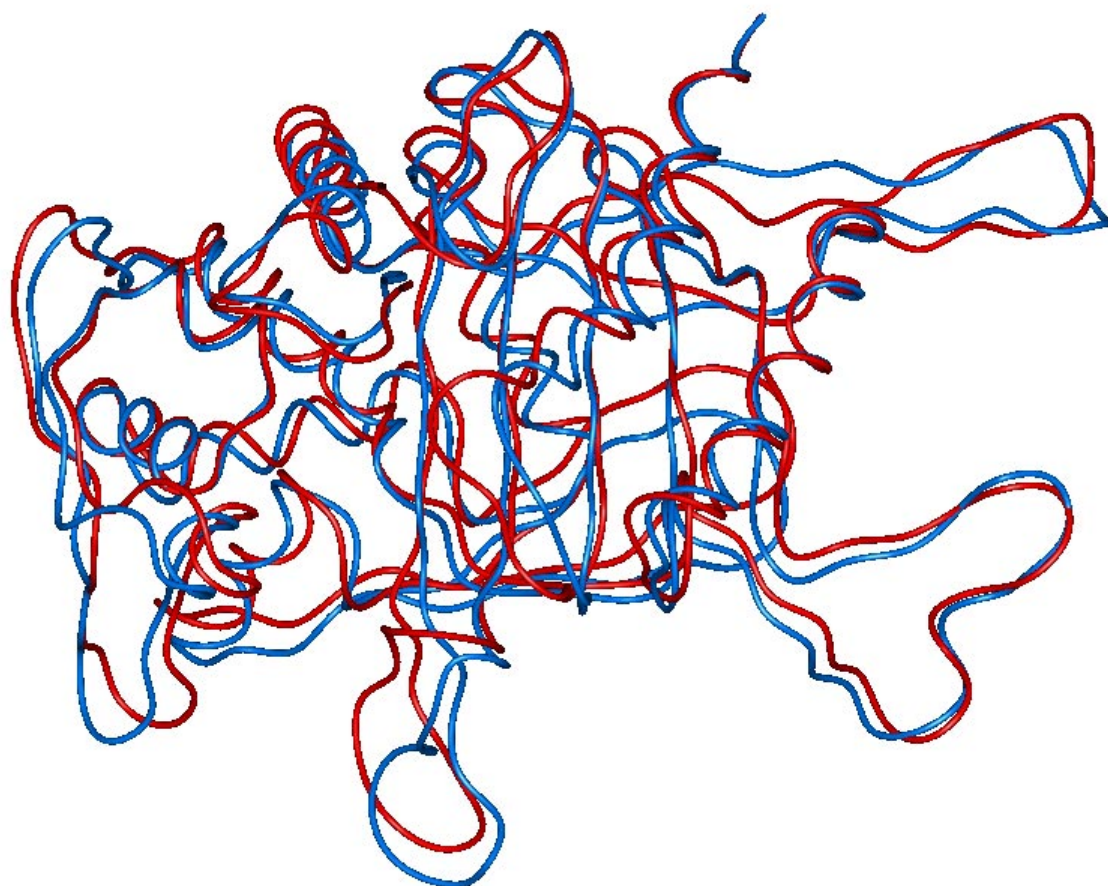
Fortunately, Dr Dideberg supplied the new X-ray coordinates of the loop spanning residues T370 to M386. So, finally, after substituting these new coordinates for the corresponding amino acids, and folding the rest of the structure according to 1PMD.pdb, the last six MD simulations were performed again. This time also the heavy atoms of T370 to M386 were restrained, in order to maintain their X-ray coordinates as much as possible. In the last simulation one of the hydrogens of the positively charged amine group of K547 and the carbonyl oxygen atom of S548 were constrained to each other using a force of 1-2 kcal/mole/Å<sup>2</sup>. Table 4.5 summarises the values of the most important parameters obtained during the second round of MD calculations. A complete analysis of these calculations is given in Appendix H.

**Table 4.5** Analysis of the various MD simulations of the penicillin-binding domain of PBP2x\* after incorporation of the new X-ray coordinates.

Temperature (K)	restrain / tether	force	simulation time (ps)	internal energy (kcal)	RMS trace (full trans- peptidase domain)	RMS trace (aligned regions)	%most favoured regions	%dis- allowed regions		distorted	
									main chain bonds	main chain angles	planar groups
310	restrain	1.0 - 2.0	220	2372.9	1.04	0.83	77.0	0.7	5	0	4
310	restrain	0.8 - 1.6	220	2354.7	1.17	0.88	80.9	0.7	6	0	3
310	restrain	0.6 - 1.2	220	2331.4	1.26	0.89	82.6	0.7	6	0	2
310	restrain	0.4 - 0.8	220	2282.5	1.39	0.90	80.3	0.7	6	0	3
310	restrain	0.2 - 0.4	220	2290.0	1.50	0.92	79.3	1.0	8	0	0
310	restrain	0	220	2282.8	1.56	0.98	79.3	1.0	7	0	1

### 4.3 The final model of *S. pneumoniae* PBP2x

Figure 4.1 shows the relative displacement of the ribbon of the transpeptidase domain of the final model of *S. pneumoniae* PBP2x with respect to the crystal structure. Both the N- and C-terminal domains have been superimposed.



**Figure 4.1** Ribbon presentation of the superposition of the  $C^{\alpha}$ -atoms of the transpeptidase domain (Ser266 to Asp616) of the crystal structure (1PMD.pdb, coloured blue) and the optimised structure of the model (coloured red). The  $C^{\alpha}$ -atoms of the N-terminal and C-terminal domains have been superimposed in order to show the relative displacement of the transpeptidase domain.

The Ramachandran plot (Figure 4.2 shown on page 93) shows that 79.3 % of the non-glycine and non-proline residues occur in the most favoured regions. According to the parameters listed in Table 4.6 this is much better than could be expected from a 3.5 Å resolution structure. Moreover, although the peptide bond planarity ( $\omega$ -angle) is with a standard deviation value of  $9.7^{\circ}$  from planarity higher than the experimentally observed value of  $5.8^{\circ}$  in X-ray structures [185], the overall G-factor is

with a value of  $-0.5$  much better than the typical value of  $1.0$ . On the other hand, a study [186] comparing 43 analogous enzymes showed that in 55% of the models the  $\omega$ -angles differed more than  $10^\circ$  from the ideal value (in 14% of the models even more than  $20^\circ$ ). Therefore, it is likely that the bad value of this parameter has resulted from the minimisation routine used in the force field.

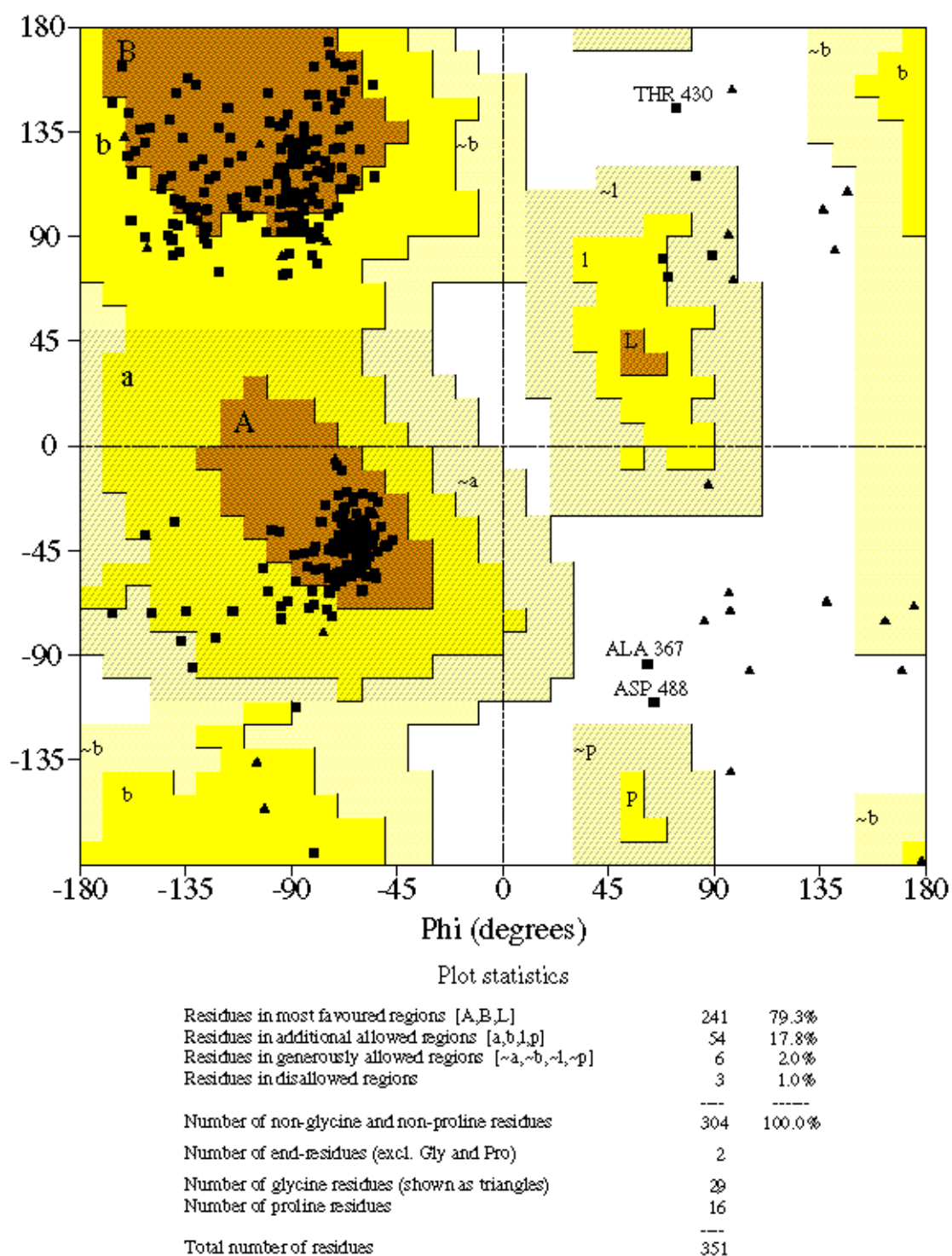
**Table 4.6** *Stereochemical parameters of the backbone of the final model of the transpeptidase domain of PBP2x.*

Stereochemical parameter	Number of data points	Parameter value	Typical value	Band width	Number of band widths from mean
% in most favoured regions	304	79.3	53.6	10.0	2.6
$\omega$ -angle standard deviation	349	9.7	6.0	3.0	1.2
bad contacts per 100 residues	0	0.0	32.6	10.0	-3.3
$\xi$ -angle standard deviation	322	2.0	3.1	1.6	-0.7
H-bond energy standard deviation	218	0.6	1.2	0.2	-2.7
Overall G-factor	351	-0.5	-1.0	0.3	1.5

The excellent quality of the final model is further shown by the fact that the stereochemical data of the side chains are much better than expected for a  $3.5 \text{ \AA}$  resolution structure. These data are listed in Table 4.7.

**Table 4.7** *Stereochemical parameters of the side chains of the final model of the transpeptidase domain of PBP2x.*

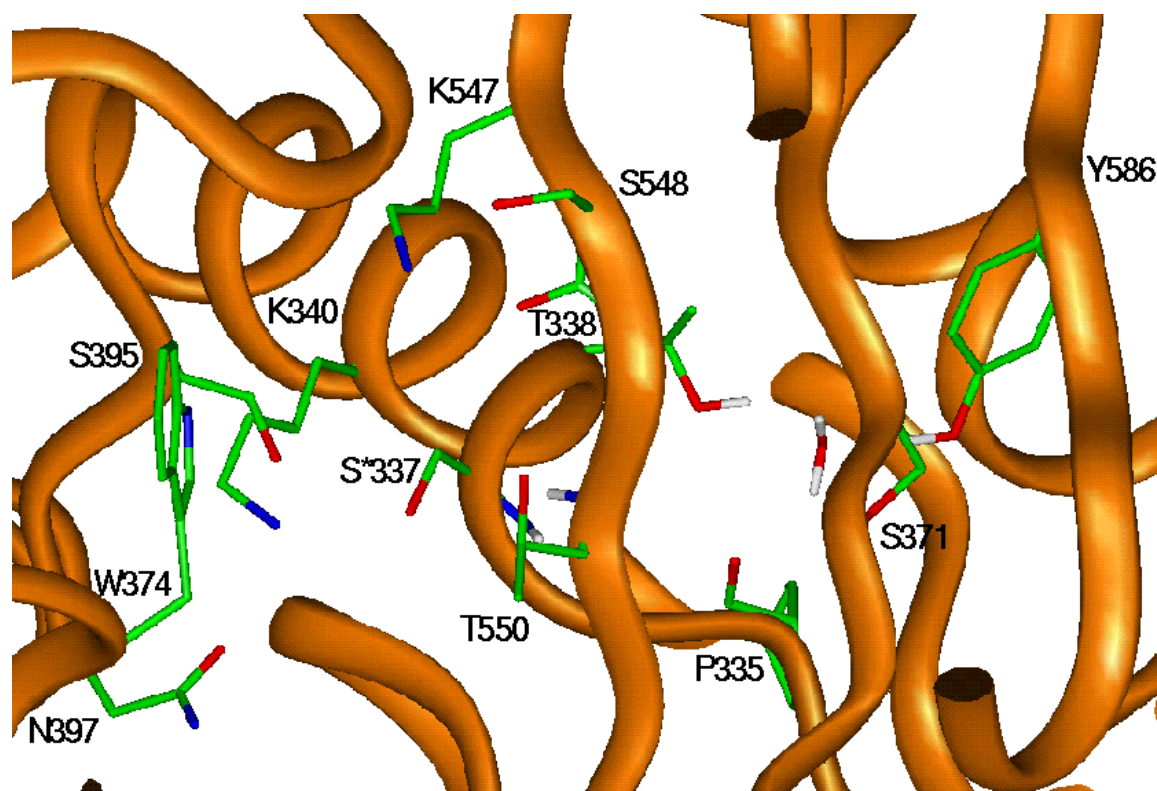
Stereochemical parameter	Number of data points	Parameter value	Typical value	Band width	Number of band widths from mean
Chi-1 gauche minus standard dev.	33	6.5	31.8	6.5	-3.9
Chi-1 trans standard deviation	125	8.5	30.1	5.3	-4.1
Chi-1 gauche plus standard dev	118	9.6	28.8	4.9	-3.9
Chi-1 pooled standard deviation	276	9.6	29.7	4.8	-4.1
Chi-2 trans standard deviation	73	10.7	28.5	5.0	-3.6



**Figure 4.2** Ramachandran plot of the optimised structure of the transpeptidase domain of PBP2x\*.

Figure 4.3 shows a view of the active site. Finally, Figure 4.4 presents a view of the active site cavity as generated with SURFNET [187].

a



b

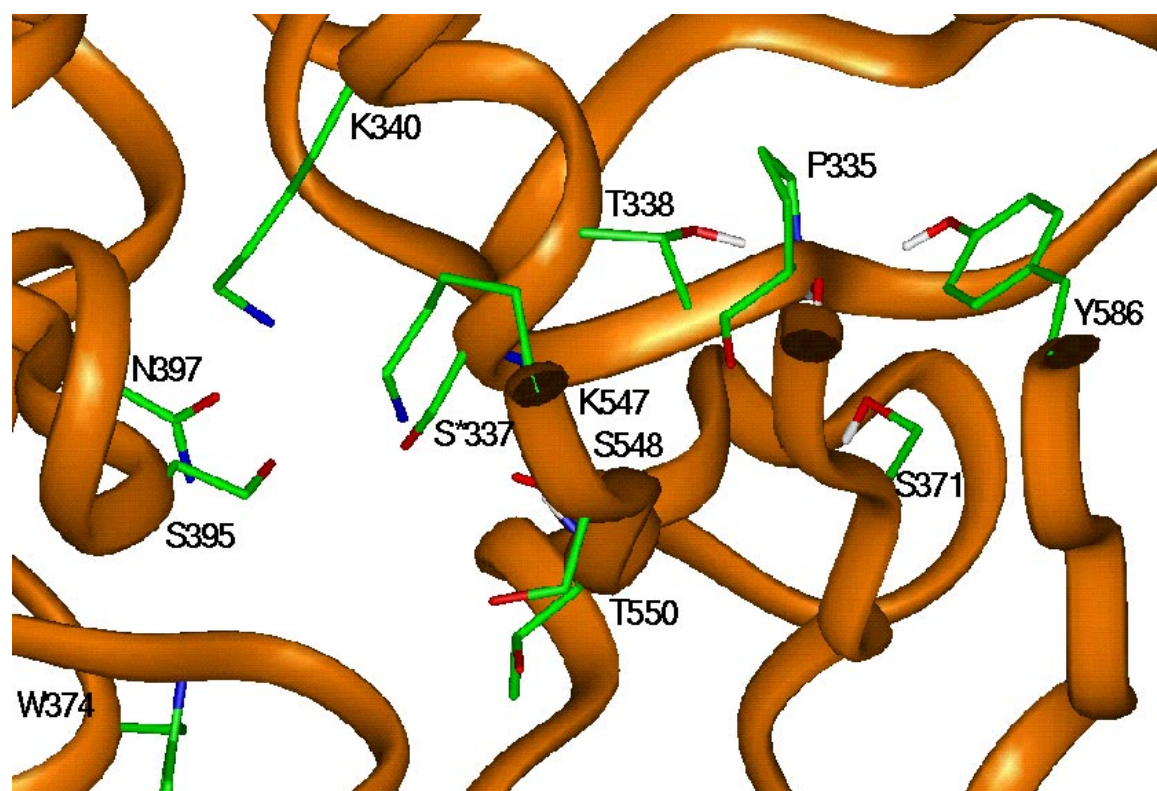
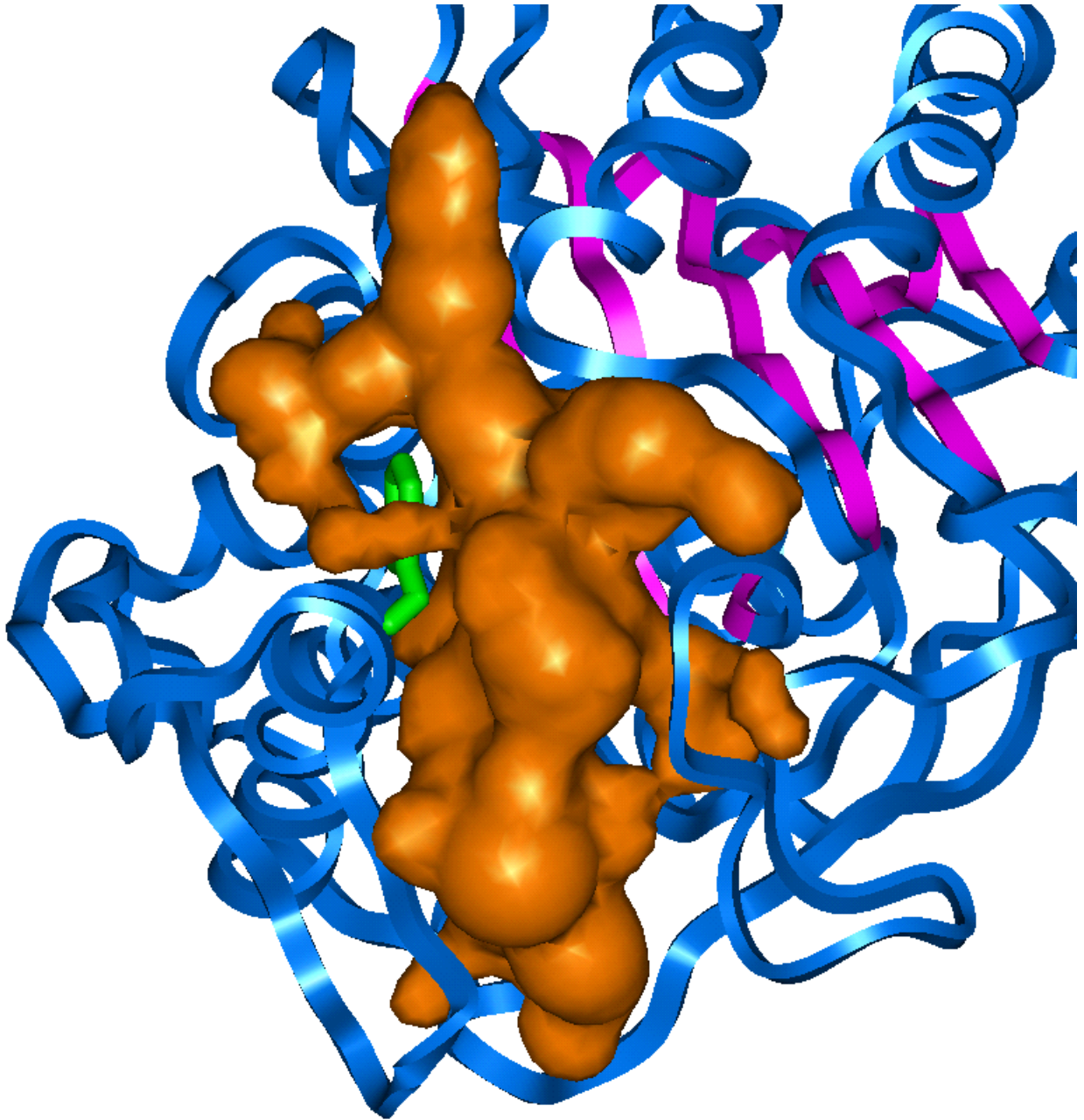


Figure 4.3 View of the active site of PBP2x\*. (a) Frontal view. (b) Top view.



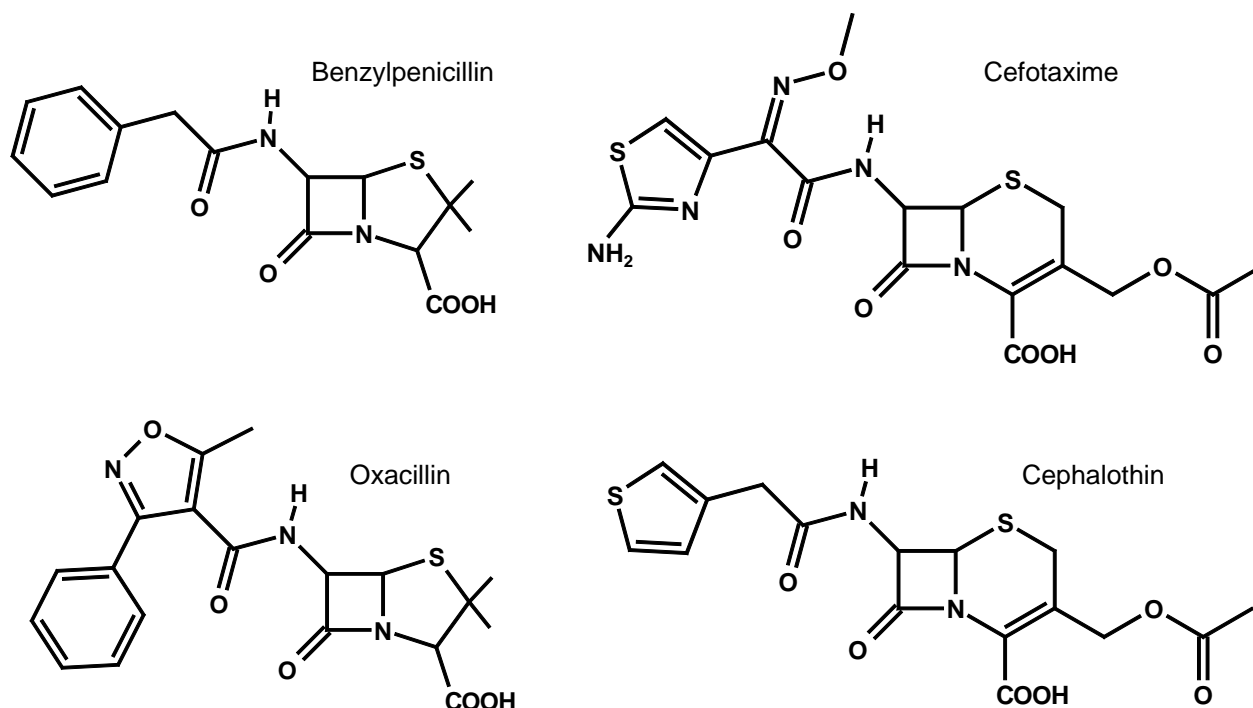
**Figure 4.4** *View of the active site cavity generated with SURFNET [187]. The five-stranded  $\beta$ -sheet is shown in magenta. The side chain of W374 is shown as sticks.*

## 4.4 Inhibitors

To further investigate the three-dimensional aspects of the binding of antibiotics by PBP2x, the active site of the transpeptidase domain was complexed with ligands, after which each of these complexes was subjected to a molecular dynamics calculation. In order to be able to accurately describe the lactam ring contained in both penicillins and cephalosporins the parameter list of the CVFF had to be expanded. The parameters needed for describing these structural parts have previously been derived [188, 189] in the group of Professor Höltje. A list of these extra parameters added to the force field is given in Appendix E.

### 4.4.1 Choice of inhibitors

The four antibiotics used as inhibitors in this project are shown in Figure 4.1.



**Figure 4.1** The four antibiotics used as inhibitors in this study.



As mentioned in section 1.6.3 the affinity of various penicillins and cephalosporins is affected upon various amino acid mutations in or near the active site. One of the most important is the T550→A mutation, since the affinity of the important third-generation cephalosporin cefotaxime is substantially lowered, whereas the affinity for benzylpenicillin remains practically unaltered (see Table 1.1). In order to detect possible differences in their binding modes, these two antibiotics were chosen as ligands.

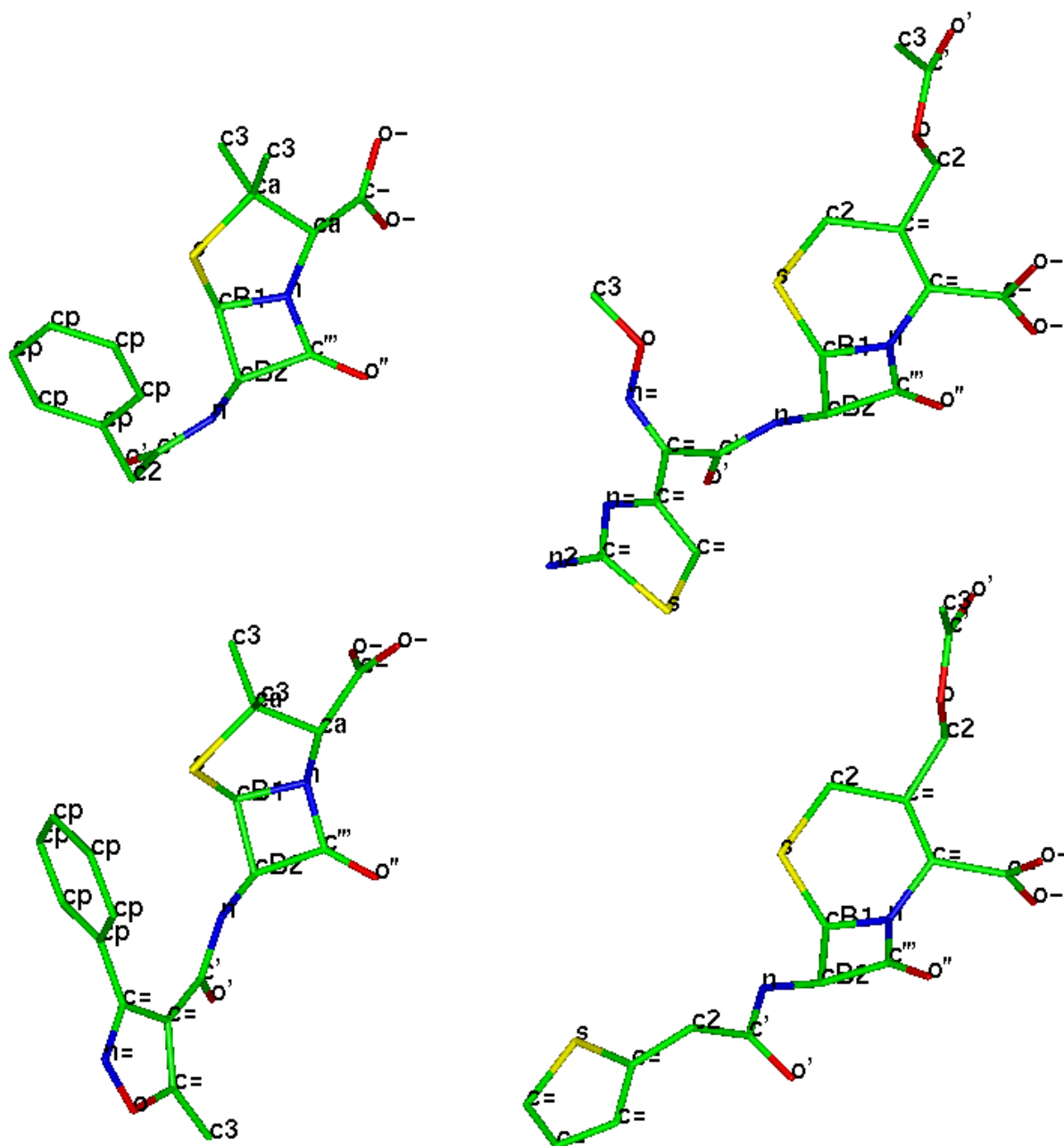
In order to investigate whether their differences in binding modes could be generalised for the penicillins and cephalosporins respectively, the penicillinase-stable penicillin oxacillin and the first-generation cephalosporin cephalothin were chosen as well. Oxacillin was also included in the studies to find a possible explanation for the fact that contrasting to the susceptibility for cefotaxime, the T550→A mutation results in an *increased* susceptibility to oxacillin [61].

Cephalothin, on the other hand, was included into the dynamics studies to verify recent IR-measurements [79] focussing on the hydrogen bonding of the carbonyl oxygen, which have shown that this cephalosporin was less stabilised in the oxyanion hole than benzylpenicillin.

#### 4.4.2 Construction of the ligands

For both benzylpenicillin and cephalothin X-ray structures were available in the CSD (reference codes NABZ and CETHNA, respectively). Both the structures of cefotaxime and oxacillin, had to be constructed taking other structures as templates, however. Thus, using the molecular modelling package SYBYL (version 6.4) [190], cefotaxime was constructed from ceftizoxime (reference code FAJMAE), by adding the C3 substituent in the conformation as present in the crystal structure of cephalothin. Likewise, oxacillin was build from the X-ray structure of benzylpenicillin by changing the C6 substituent. For each antibiotic, the atom types as shown in Figure 4.1 were added. Because of parameterisation reasons, however, the charges on the sulfur atoms present in both the penicillins and cephalosporins on the one hand, and on the atoms constituting the thiazole ringsystem in cefotaxime on the other hand, would not have correctly been computed in the force field. Therefore,

before placing them in the enzyme active site, each of these structures was semiempirically optimised using the AM1 Hamiltonian.



**Figure 4.1** Atom types of the four antibiotics used as inhibitors in the complexation studies. For explanation see Appendix E.

## 4.5 Enzyme-inhibitor complexes

As explained in section 1.4.2 the covalent binding and eventual release of  $\beta$ -lactam antibiotics by active site serine proteins is a complex three-step process. Since bonds are formed and broken during the course of this process, detailed study of the whole process with force field methods is impossible. Study of the formation of a Michaelis complex is possible at the most with these methods. However, although the acylation and deacylation steps can not be studied in detail, thorough investigation of the three-dimensional aspects of these complexes should give at least some hints as to why some antibiotics are more potent than others.

Having treated both the enzyme and the individual inhibitors solely up to now, in the following low-energy and sensible protein-inhibitor complexes will be built. The stability of these complexes will be shown with molecular dynamics simulations.

### 4.5.1 Docking inhibitors into the active site of PBP2x

Both benzylpenicillin and cefotaxime were first manually docked into the active site of the final model of transpeptidase domain of PBP2x generated in section 4.2.4. In order to see whether any better complex conformations could be found, these conformations were taken as input for AutoDock 2.4 [173]. Lone pairs were added as well.

The docking procedures were performed by successive action of the scripts presented in Appendix I. First, pbpq files were generated for both the inhibitor and the receptor. After generation of these files a grid surrounding the receptor was calculated. Concomitantly, the interaction energies of each atom type present in the inhibitor (charge  $+e$ ) with the receptor were stored on each grid point. Abnormally high values at grid points very close to the receptor were removed.

Next, inhibitor flexibility was specified. All bonds that were not contained in ring systems were allowed to rotate, except for the peptide bonds and the methyl groups attached to the thiazolidine ring system present in the penicillins.

After completion of this preparatory work, docking was started. One docking job consisted of 100 runs, each consisting of 50 constant temperature cycles. Starting from a random position of the inhibitor position for the initial run, new coordinates were generated by randomly changing the current position, orientation and conformation. Based on the probability of acceptance (equation 3.23), this process was repeated until 30,000 steps had been either accepted or rejected. Full configurational space was explored during the job by using translation, quaternion and torsion rotation steps of 0.2 Å, 5 degrees and 5 degrees, respectively. Each new cycle was started with the minimum conformation found in the previous conformation. In order to avoid simulated quenching, each of these 50 cycles was performed at 0.95 times the annealing temperature of the previous run, starting at 500 RT. Based on an RMS difference of 1.0, the resulting families of minimum conformations found during the job were ranked in order of increasing energy.

Because of full exploration of the configurational space, many configuration families were found, most of which representing nonrealistic solutions to the docking problem. After visual inspection of lowest configuration of each family, though, it appeared that often only one or two configurations were realistic. For both complexes the energetically most favourable conformation was selected as starting configuration for MD simulations with DISCOVER. Starting conformations of the complexes with oxacillin and cephalothin were generated by fitting the thiazolidine and dihydrothiazine ring systems, respectively. In addition, the oxime substituent of cefotaxime was positioned in the same orientation as in the covalently bound complex with cefuroxime visualised in the article by Mouz et al. [78].

## **4.5.2 Molecular dynamics simulations of PBP2<sub>xR6</sub> complexed with inhibitors**

### **4.5.2.1 Simulation conditions**

Because of the use of a different force field (AMBER) for the docking simulations in AutoDock, each selected configuration had to be minimised in the (modified) CVFF force field before starting the MD simulations. Each complex was subjected to a 240

ps MD simulation with Discover (Version 97.0). The simulation was initialised at 110 K during the first 10 ps, after which the system was simulated for another 10 ps at the same temperature. During the next 20 ps each system was heated from 110 to 310 K, after which the system was simulated for 200 ps at 310 K.

In each simulation both the inhibitor and the structurally conserved water molecule behind B3 were completely free to move. However, at the time the simulations of the complexes were started, the position of another structural water molecule became known [82]. This second structural water molecule appears to be located in the active site between K547 and the opposite helix, forming hydrogen bonds with the oxygen atoms of the backbone carbonyl groups of F392 and A393, and with the side chain hydroxyl group of T526. The water molecule is additionally stabilised by a salt bridge with the charged amine group of K547. Nevertheless, the hydrogen bonds with the backbone carbonyl groups had to be restrained with forces of 1.0 – 2.0 kcal/mole/Å each. The regions in which C<sup>α</sup> atoms were additionally positionally restrained are listed in Table 4. 8.

**Table 4. 8** *Amino acid regions of PBP2x in which the C<sup>α</sup>-atoms have been constantly restrained with 1.0 – 2.0 kcal/mole/Å during the molecular dynamics simulations of the complexes.*

Region	Restraining force (kcal/mole/Å)	Region	Restraining force (kcal/mole/Å)
F310 – Y333	1.0 – 2.0	V516 – P540	1.0 – 2.0
T352 – F392	10 – 20	G541 – F570	10 – 20
F419 – D440	1.0 – 2.0	Q590 – D616	1.0 – 2.0

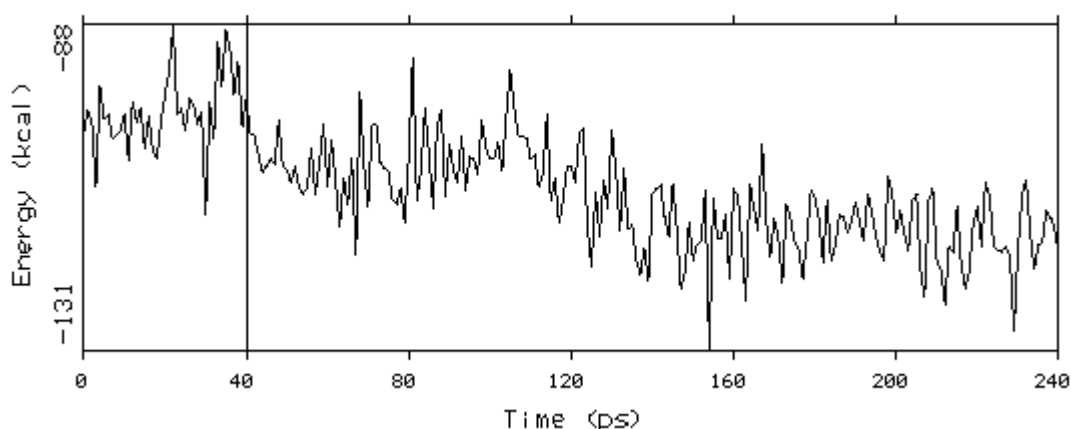
It was found that the calculations only run, if the cutoff distance was considerably increased to 33.0 – 34.5 Å.

In order to be able to carefully analyse the calculations afterwards, the intermolecular interaction energy was monitored during the simulations. Since it was found that the van der Waals energy contribution (Equation 3.38 on page 72) to the intermolecular interaction energy between the respective inhibitors and the protein was much larger than the electrostatic contribution (Equation 3.22 on page 50), whereas electrostatic

forces are likely to dominate the formation of a Michaelis complex, the dielectric constant was reduced to 1.0r.

#### 4.5.2.2 Simulation of PBP2x\*<sub>R6</sub> complexed with benzylpenicillin

Although it is well known that energy evaluations derived from force field calculations do not represent the whole truth, comparison of these data should give clues as to the relative stability of the complexes calculated. A detailed analysis of the interaction energies of benzylpenicillin complexed to PBP2x\* is given by the graphs shown in Figures 4.13 to 4.17. As follows from Figure 4.1 the interaction energy between the protein and the ligand decreases after physiological temperature has been reached at 40 ps, and decreases even further after about 100 ps simulation. This energy decrease is explained by the distance graphs presented in Figure 4.7 shown on page 105, which clearly show that a stable complex is formed in the second half of the simulation. Nevertheless, the course of the RMSD value of the inhibitor presented in Figure 4.6 (page 104) shows that it remains in the active site throughout the

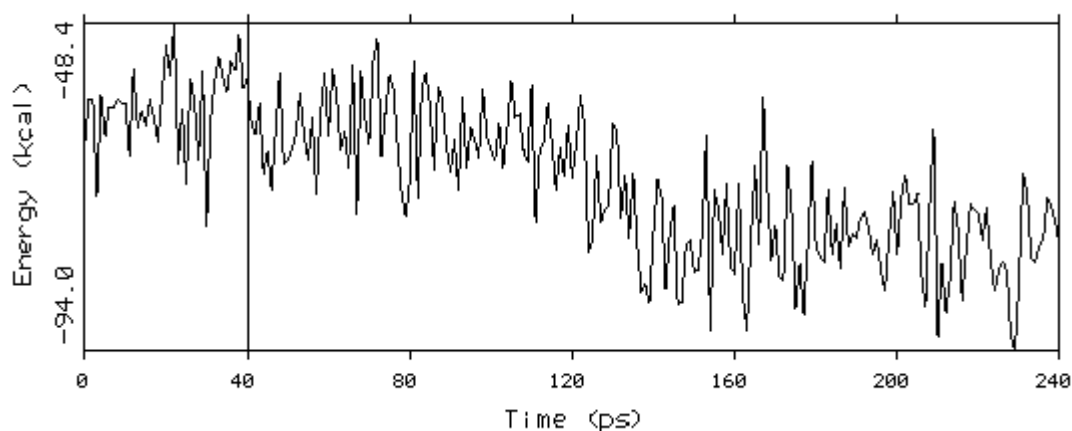


simulation.

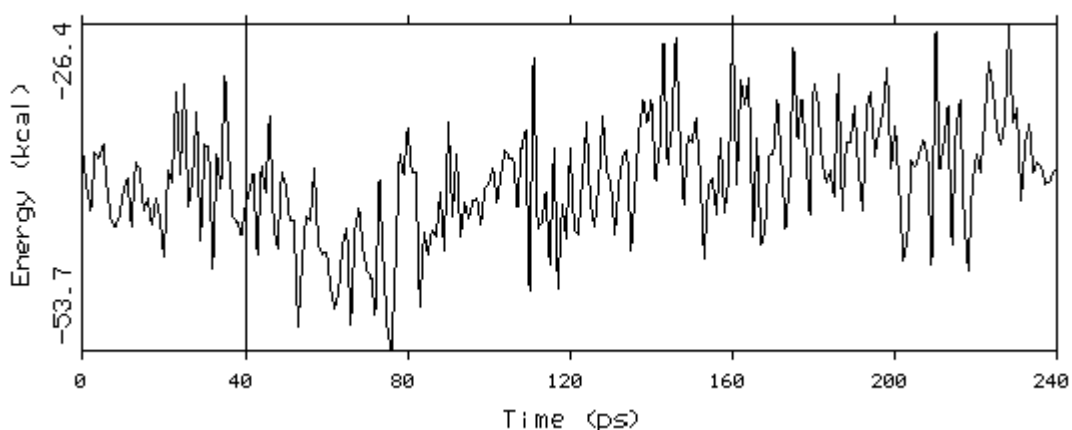
**Figure 4.1** Course of the total interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and benzylpenicillin during 240 ps MD simulation.

When both the electrostatic and Van der Waals contributions to the interaction energy are viewed separately (graphs shown in Figures 4.14 and 4.15, respectively), it is evident that this decrease is caused by a significant increase of the electrostatic interaction during the second half of the simulation. Due to tight complex formation

after the first half of the simulation only a slight increase of the Van der Waals energy is observed.

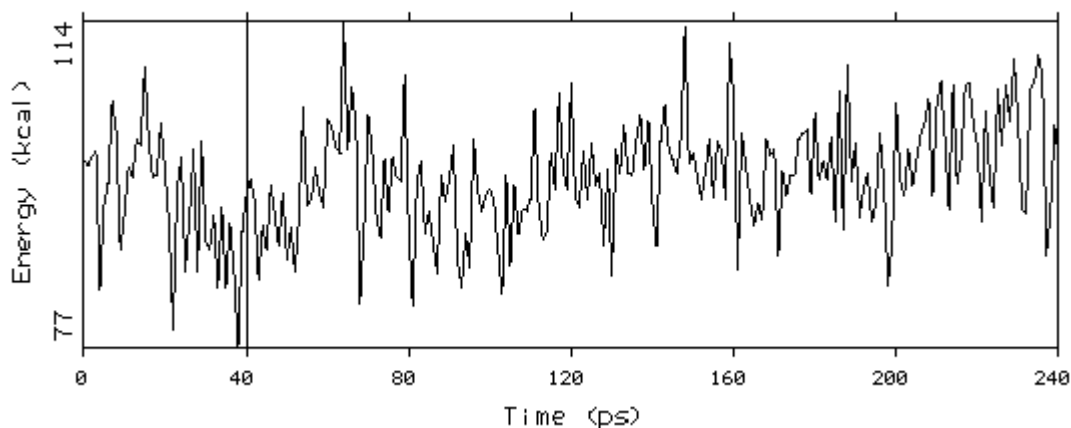


**Figure 4.2** Course of the Coulomb electrostatic contribution of the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and benzylpenicillin during 240 ps MD simulation.

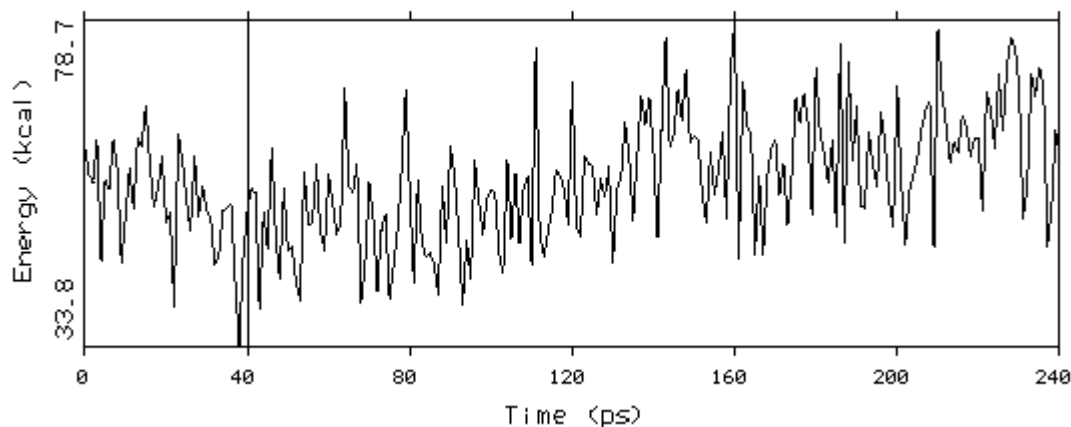


**Figure 4.3** Course of the Van der Waals contribution to the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and benzylpenicillin during 240 ps MD simulation.

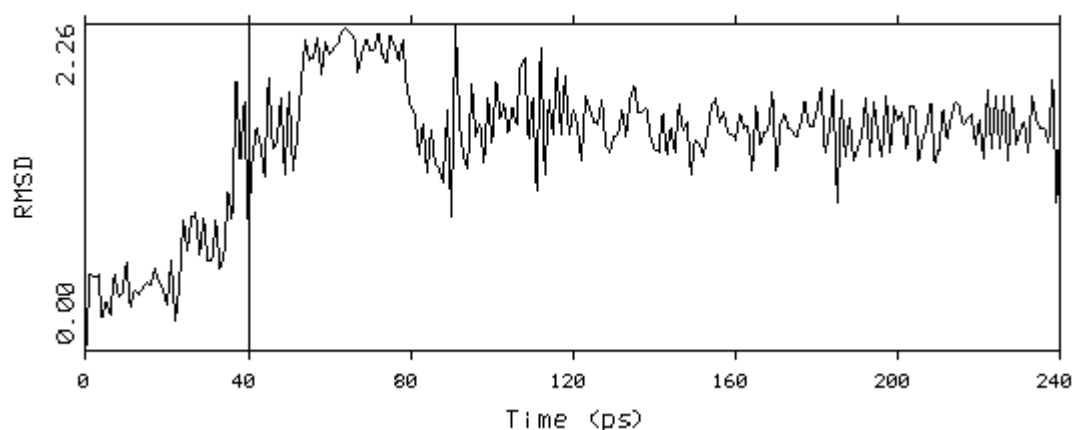
Figures 4.16 and 4.17 show the respective contributions of the dispersion and repulsion terms to the Van der Waals energy. Whereas the dispersion energy remains more or less constant throughout the simulation, the repulsion energy slightly increases after 100 ps due to tight complex formation. The stability of the complex during the second half of the simulation is further shown by the course of the RMS deviation of the antibiotic given in Figure 4.6. In addition, Figure 4.7 shows the courses of some important intermolecular distances.



**Figure 4.4** Course of the dispersion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and benzylpenicillin during 240 ps MD simulation.

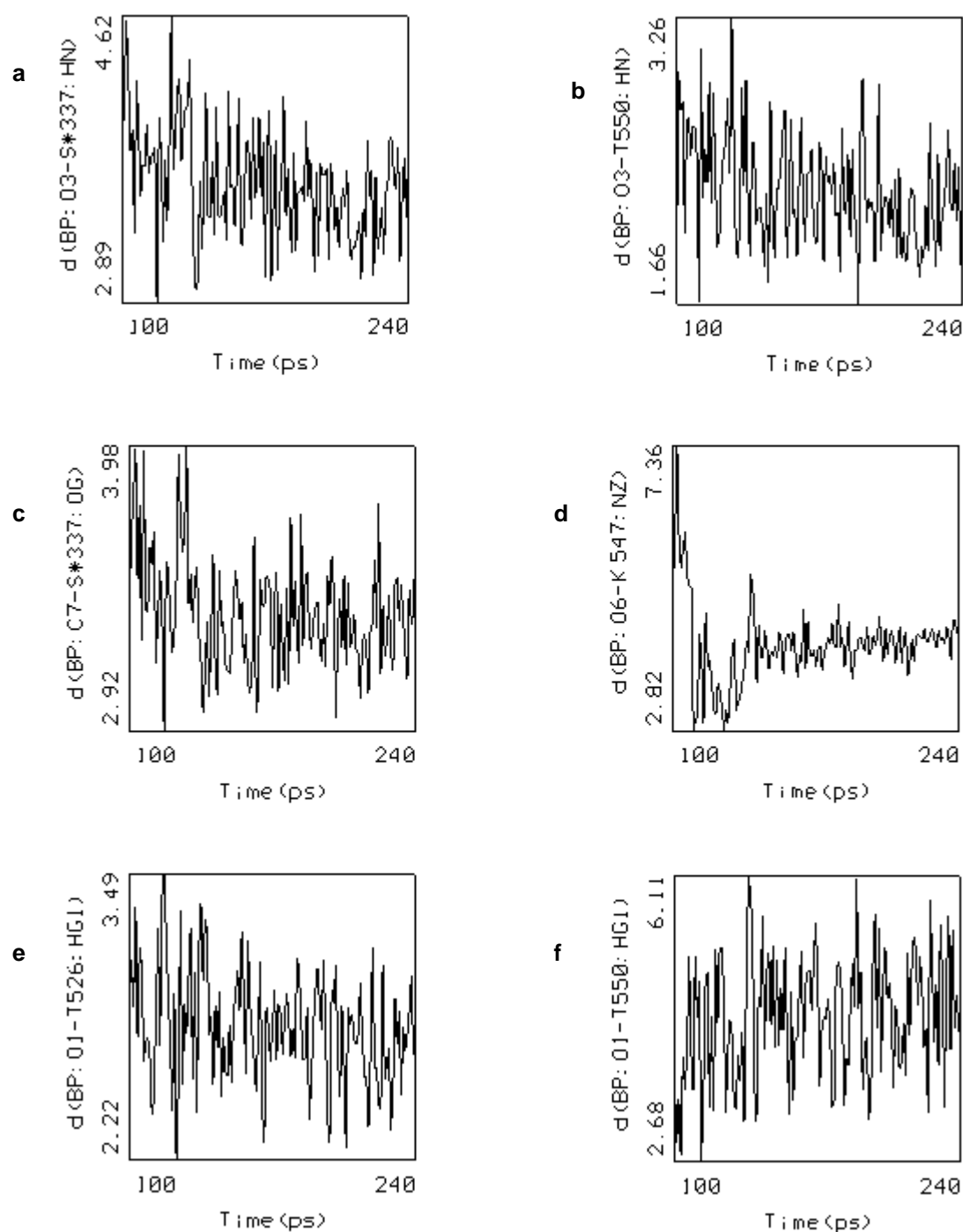


**Figure 4.5** Course of the repulsion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and benzylpenicillin during 240 ps MD simulation.



**Figure 4.6** Course of the RMS deviation of benzylpenicillin from the start configuration during the 240 ps simulation of the complex with *S. pneumoniae* PBP2x (R6 strain).





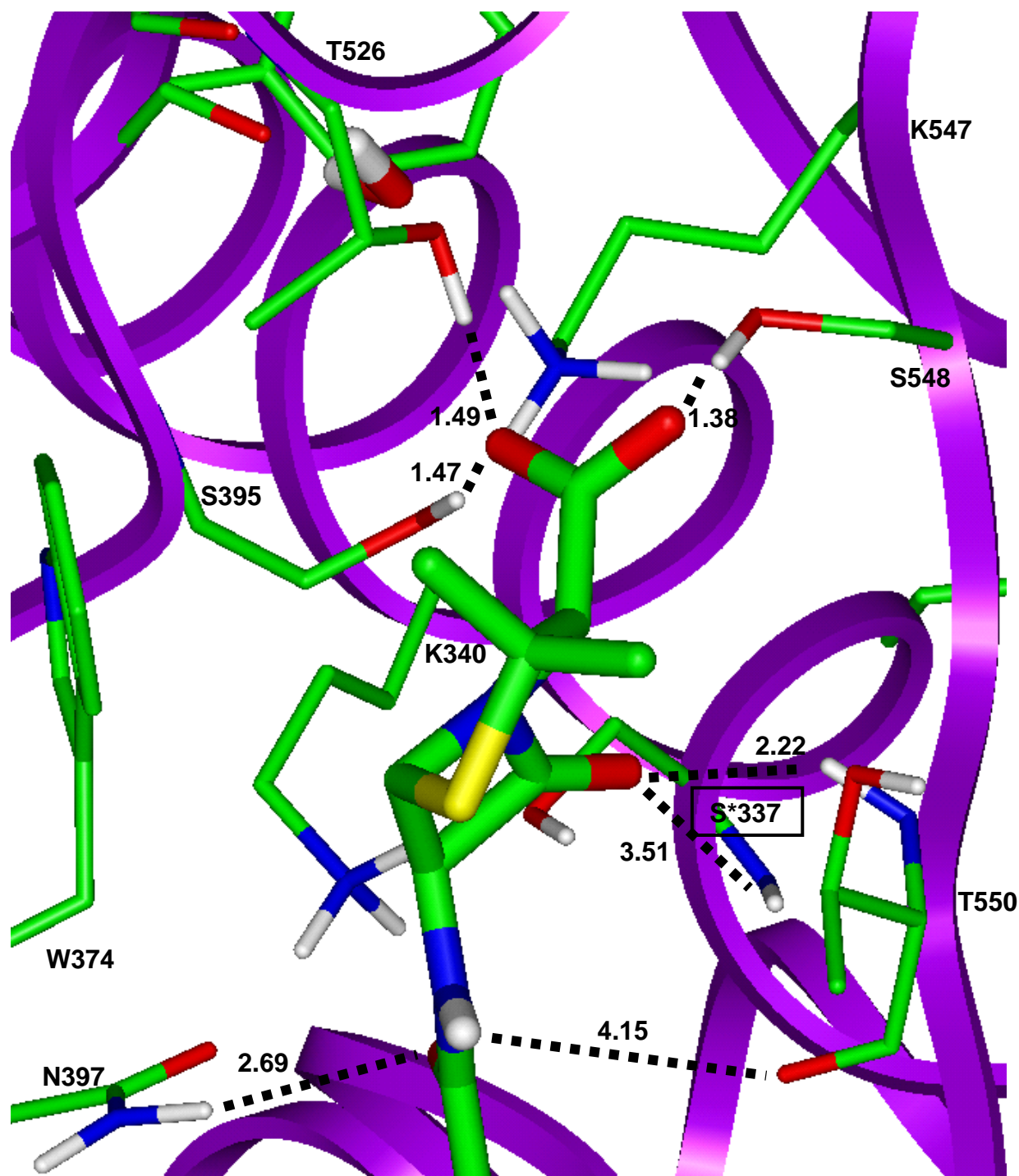
**Figure 4.7** Courses of various intermolecular atomic distances between PBP2x\* and benzylpenicillin during the last 140 ps simulation of the complex. <sup>a</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of S\*337 <sup>b</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of T550 <sup>c</sup>Distance between the electrophilic lactam carbonyl carbon atom and the nucleophilic hydroxyl oxygen of S\*337 <sup>d</sup>Distance of the salt bridge between the charged nitrogen atom of K547 and the nearest carboxylate oxygen atom <sup>e</sup>Distance of the hydrogen bond between the hydroxyl hydrogen atom of T526 and the carboxylate group <sup>f</sup>Distance of the hydrogen bond between the hydroxyl hydrogen of T550 and the carboxylate group

In order to only take the energetically more favourable second half of the simulation into account, Figure 4.8 presents the average structure of the complex during the last 140 ps simulation.



**Figure 4.8** Structure of PBP2x\* complexed with benzylpenicillin, averaged over the last 140 ps of the MD simulation. The average structure of benzylpenicillin is shown as atom type coloured sticks. The starting configuration is shown as thick black sticks, whereas the structures shown as small black sticks represent the conformation of benzylpenicillin at 120 ps, 180 ps and 240 ps.

Figure 4.9 shows the Michaelis complex (average structure from 100 to 240 ps) zoomed in on the active site.



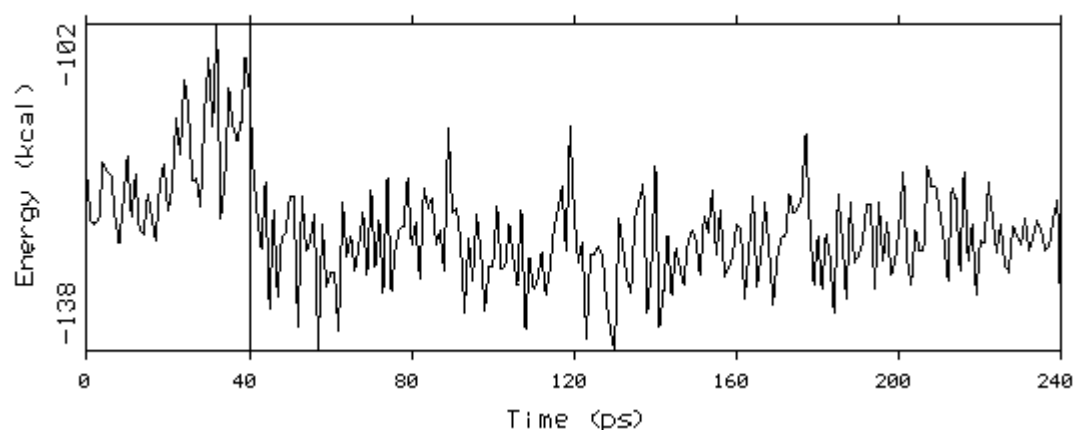
**Figure 4.9** Average structure of PBP2x\* (taken over the last 140 ps) complexed with benzylpenicillin showing the hydrogen bonding interactions in the active site (dashed lines). Intermolecular distances shown represent average distances.

Throughout the simulation, the carboxylate group of benzylpenicillin was stabilised by hydrogen bonds with the side chains of S395, T526 and S548, whereas due to  $sp^3$  hybridisation of C2 a hydrogen bond with the  $\gamma$ -OH of T550 was geometrically

hindered. The amide group of the inhibitor does not seem to be able to adopt a suitable position for optimal stabilisation by the side chain of N397 and the backbone carbonyl group of T550. Instead, the amide oxygen atom is more attracted by the positively charged N $\epsilon$  of K340.

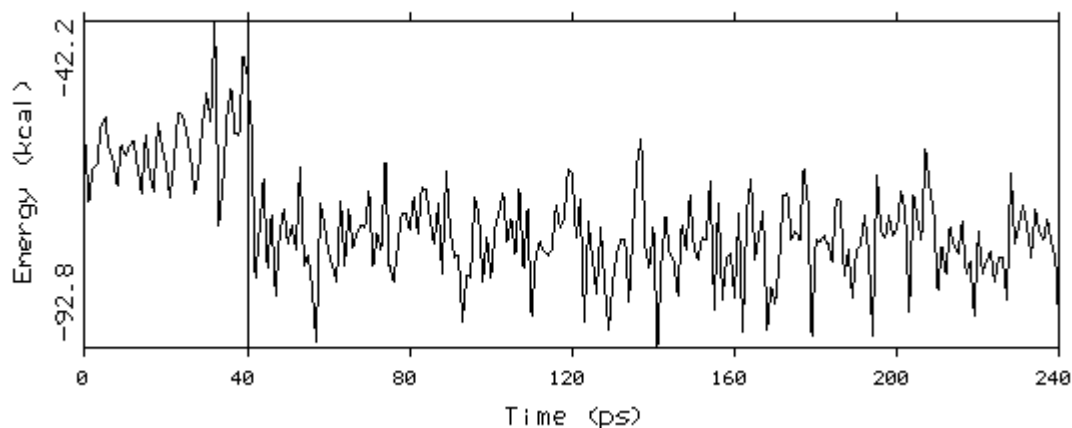
#### 4.5.2.3 Simulation of PBP2x\*<sub>R6</sub> complexed with cefotaxime

A detailed analysis of the interaction energies of cefotaxime complexed to PBP2x\* is given by the graphs shown in Figures 4.22 to 4.26. As follows from Figure 4.1 a stable configuration of the complex is adopted immediately after 40 ps simulation time, when physiological temperature has been reached.

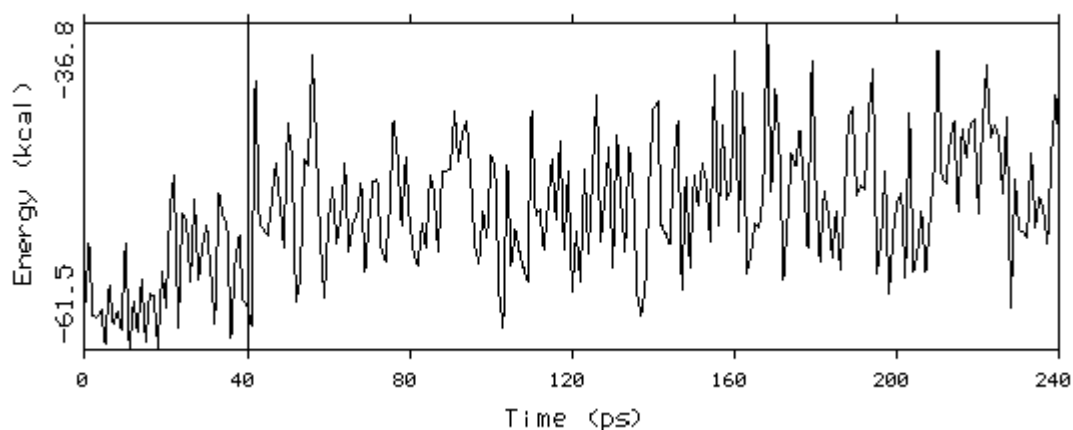


**Figure 4.1** Course of the total interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cefotaxime during 240 ps MD simulation.

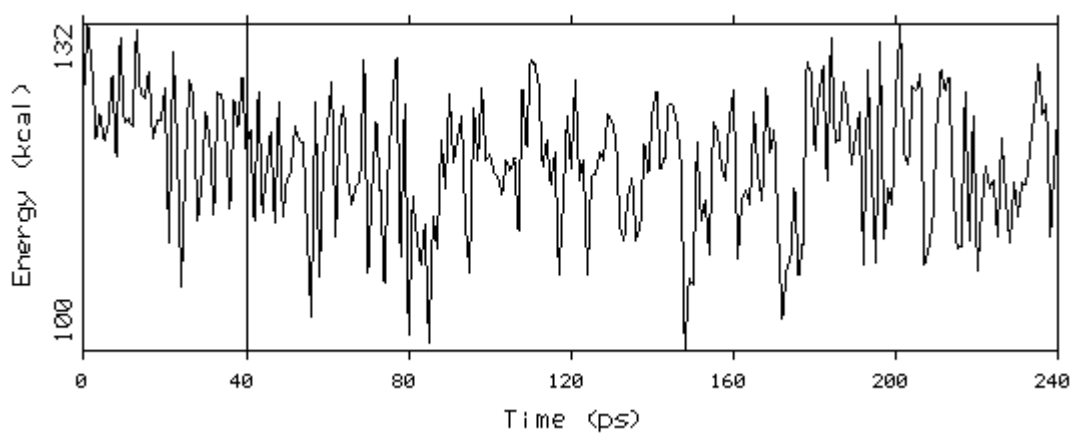
As shown in Figures 4.23 and 4.24, respectively, both the electrostatic and Van der Waals contributions to the interaction energy remain stable after the initial 40 ps. Moreover, as indicated by the graphs shown in Figures 4.25 and 4.26, both the dispersion and repulsion energies remain constant throughout the simulation. The stability of the complex is further shown by the graph presented in Figure 4.6, showing the course of the RMS deviation of cefotaxime during the simulation.



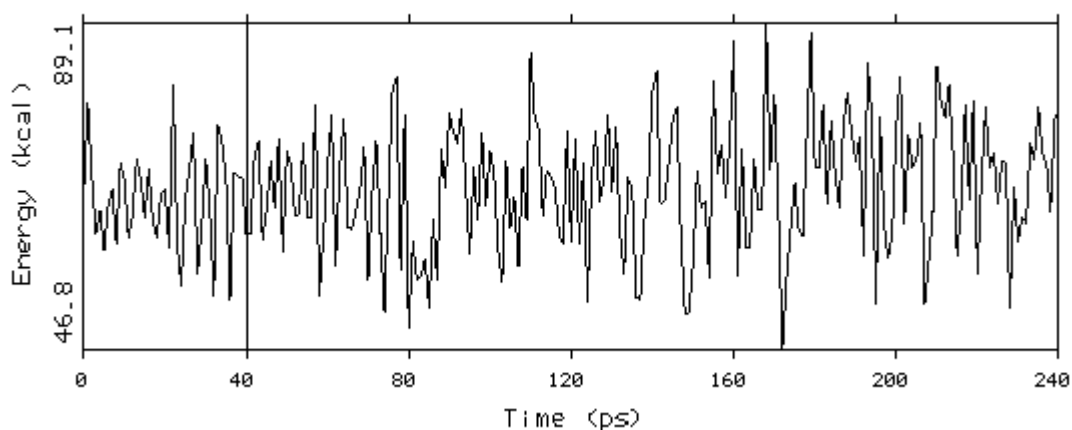
**Figure 4.2** Course of the Coulomb electrostatic contribution of the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cefotaxime during 240 ps MD simulation.



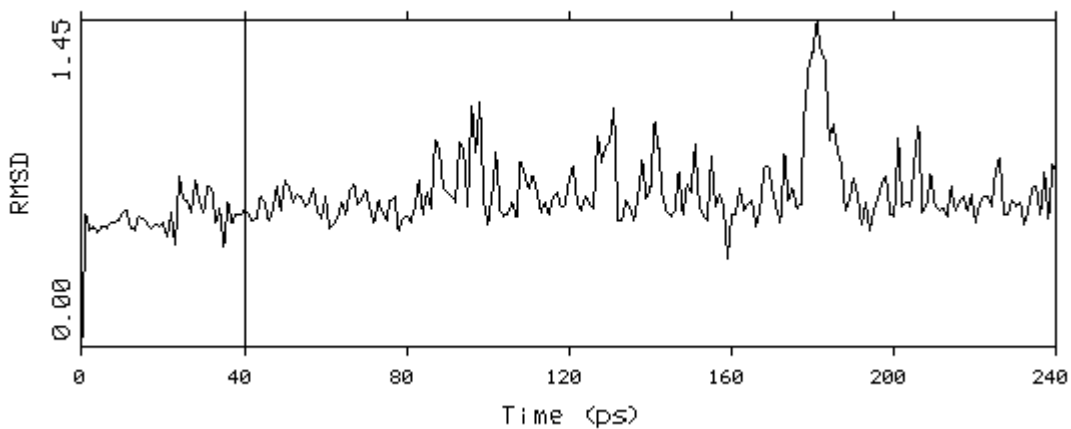
**Figure 4.3** Course of the Van der Waals contribution to the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cefotaxime during 240 ps MD simulation.



**Figure 4.4** Course of the dispersion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cefotaxime during 240 ps MD simulation.

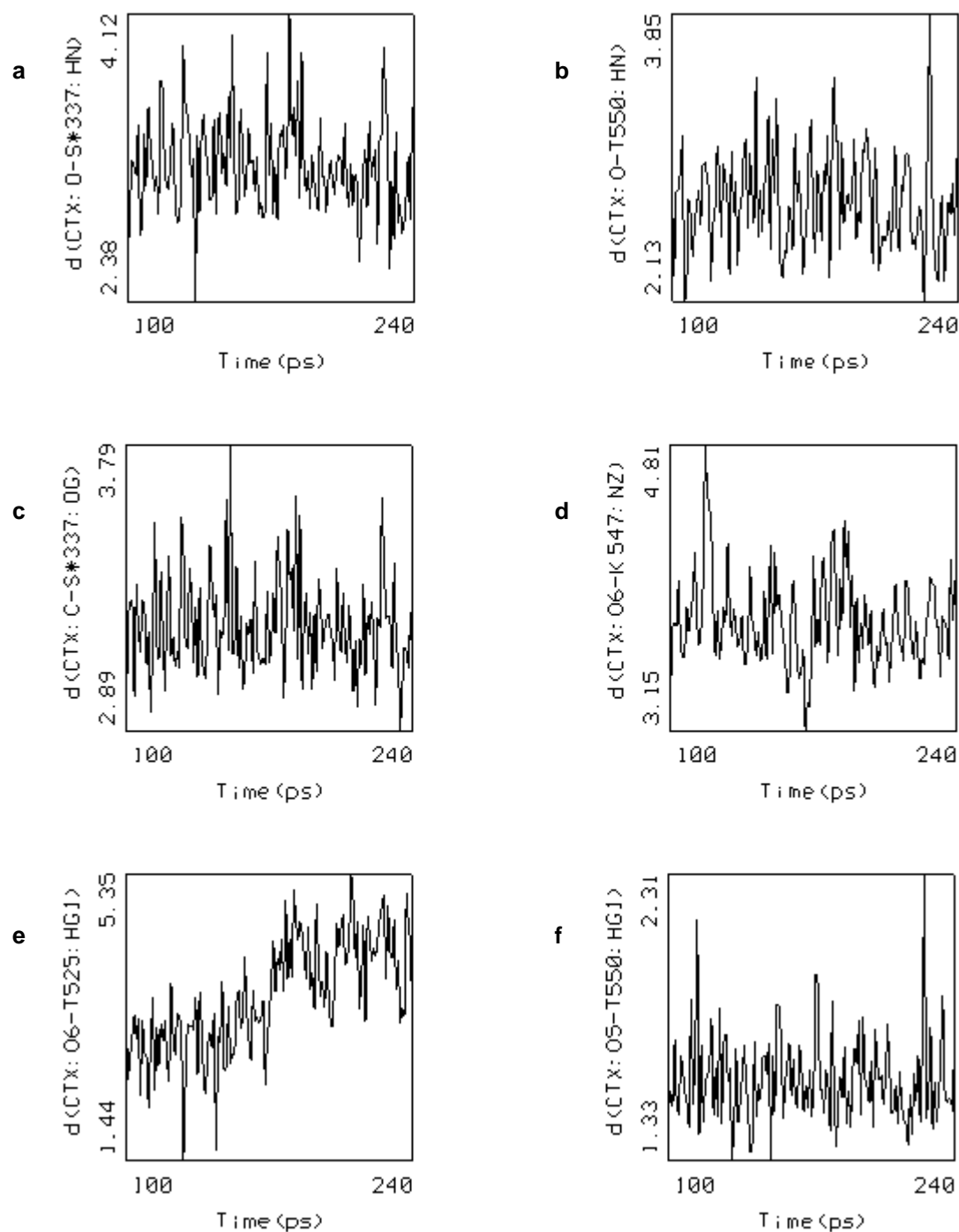


**Figure 4.5** Course of the repulsion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cefotaxime during 240 ps MD simulation.



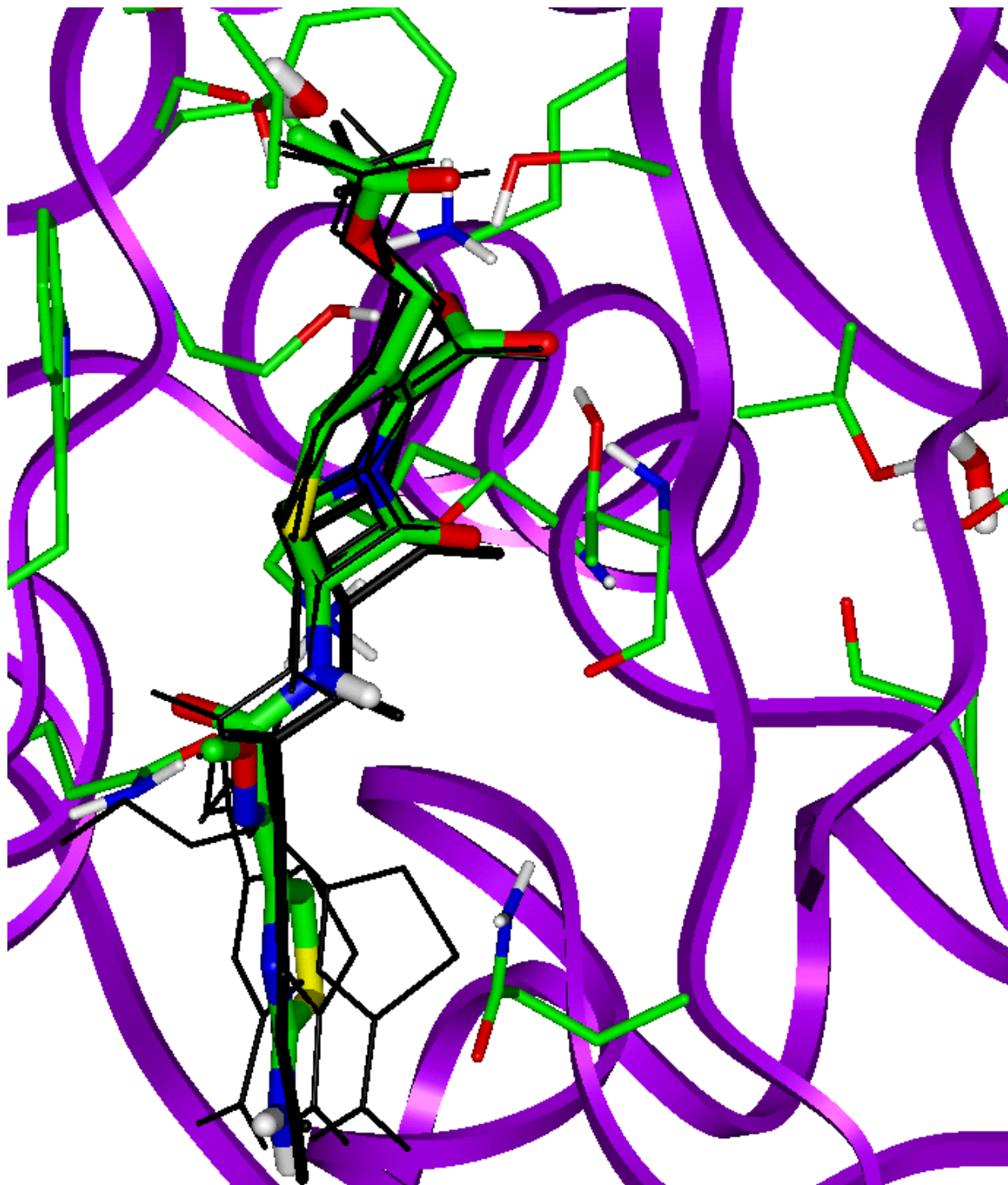
**Figure 4.6** Course of the RMS deviation of cefotaxime from the start configuration during the 240 ps simulation of the complex with *S. pneumoniae* PBP2x\* (R6 strain).

The distance graphs presented in Figure 4.7 even more confirm the stability of the complex of PBP2x\* and cefotaxime.



**Figure 4.7** Courses of various intermolecular atomic distances between PBP2x\* and cefotaxime during the last 140 ps simulation of the complex. <sup>a</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of S\*337 <sup>b</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of T550 <sup>c</sup>Distance between the electrophilic lactam carbonyl carbon atom and the nucleophilic hydroxyl oxygen of S\*337 <sup>d</sup>Distance of the salt bridge between the charged nitrogen atom of K547 and the nearest carboxylate oxygen atom <sup>e</sup>Distance of the hydrogen bond between the hydroxyl hydrogen of T526 and the carboxylate group <sup>f</sup>Distance of the hydrogen bond between the hydroxyl hydrogen atom of T550 and the carboxylate group.

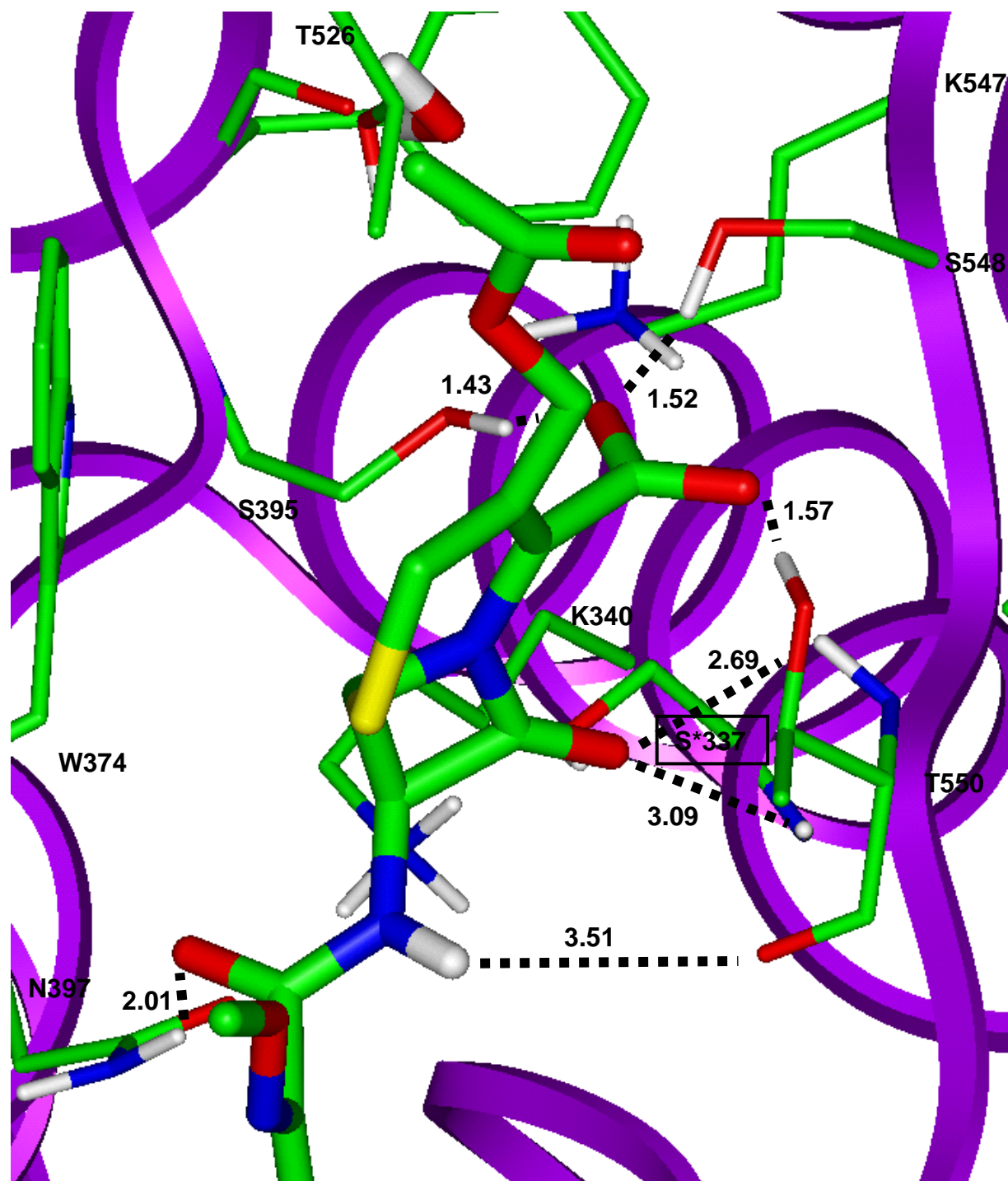
In order to be able to fully compare the average structures of the complex, the structure shown in Figure 4.8 also represents the average structure of the last 140 ps of the simulation.



**Figure 4.8** Structure of PBP2x\* complexed with cefotaxime, averaged over the last 140 ps of the MD simulation. The average structure of benzylpenicillin is shown as atom type coloured sticks. The starting configuration is shown as thick black sticks, whereas the structures shown as small black sticks represent the conformation of cefotaxime at 120 ps, 180 ps and 240 ps.



Finally, Figure 4.9 shows the Michaelis complex (average structure from 100 to 140 ps) zoomed in on the active site.



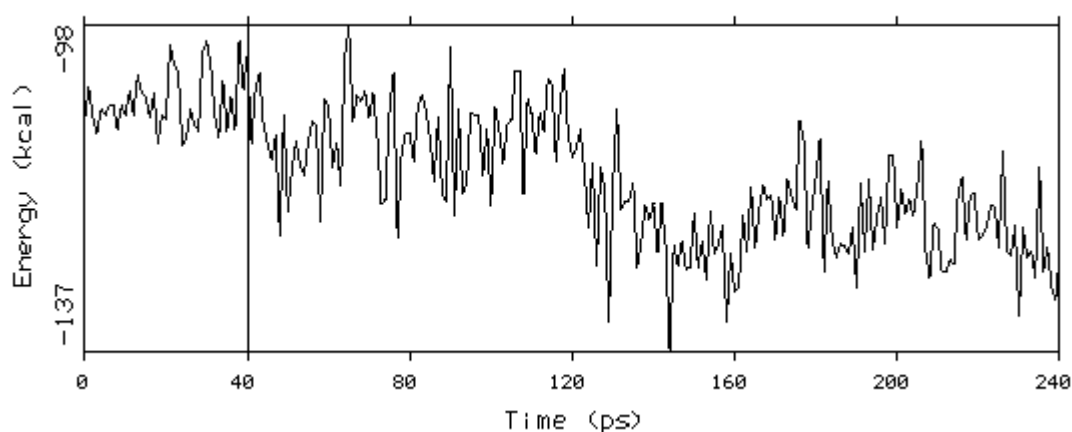
**Figure 4.9** Average structure of PBP2x\* complexed with cefotaxime showing the hydrogen bonding interactions in the active site (shown as sticks). Intermolecular distances shown represent average distances.

Unlike the complex with benzylpenicillin, a strong hydrogen bond exists between the carboxylate group and the  $\gamma$ -OH of T550. Instead, the  $\gamma$ -OH of T526 is turned away from the carboxylate group, thereby stabilising S395 and enabling hydrophobic

interaction between its methyl group and the dihydrothiazine ring system. Compared to benzylpenicillin, however, the amide group is now much stabilised.

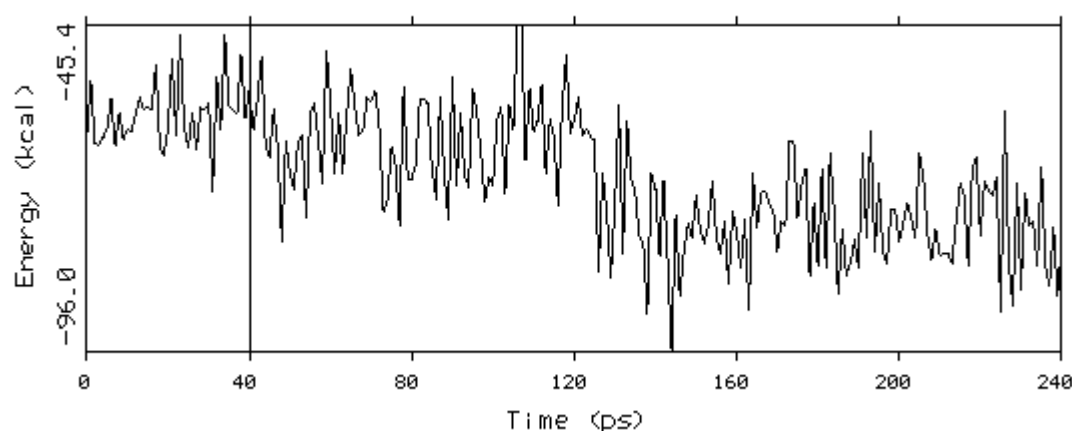
#### 4.5.2.4 Simulation of PBP2x\*<sub>R6</sub> complexed with oxacillin

A detailed analysis of the interaction energies of oxacillin complexed to PBP2x\* is given by the graphs shown in Figures 4.31 to 4.35. As follows from Figure 4.1 the interaction energy between the protein and the ligand again decreases after 40 ps, decreasing even further after about 120 ps simulation.

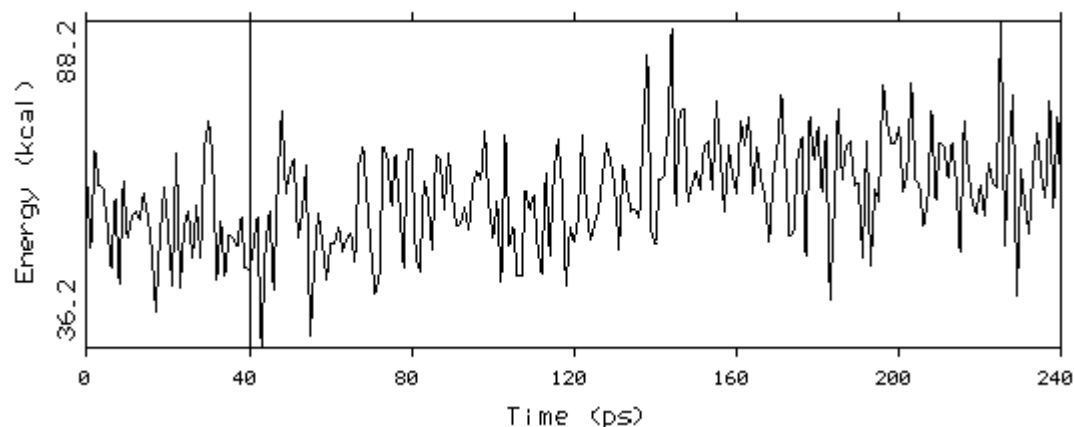


**Figure 4.1** Course of the total interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and oxacillin during 240 ps MD simulation.

Again, a significant decrease in the electrostatic interaction energy can be observed

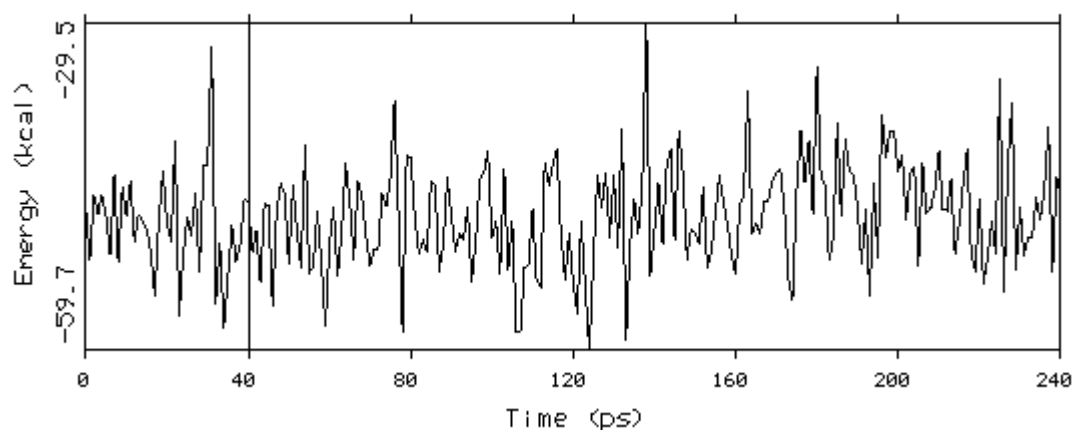


**Figure 4.2** Course of the Coulomb electrostatic contribution of the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and oxacillin during 240 ps MD simulation.

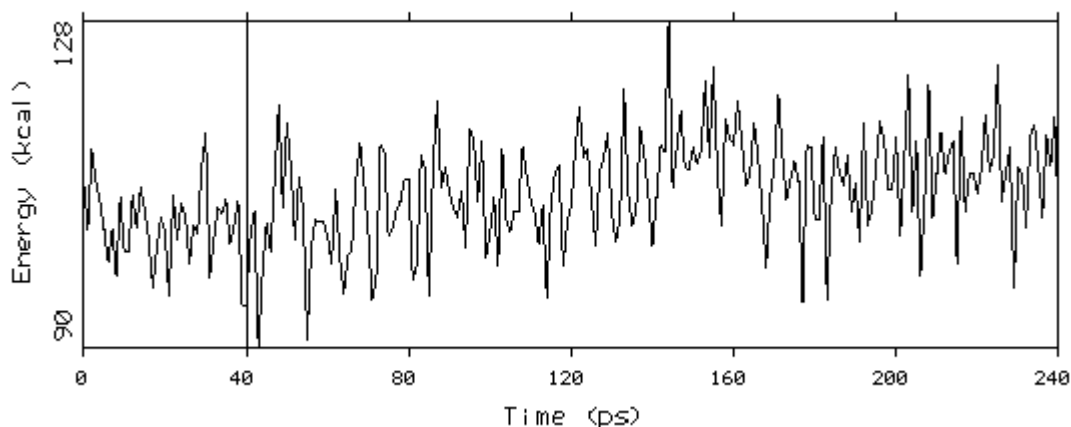


**Figure 4.3** Course of the Van der Waals contribution to the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and oxacillin during 240 ps MD simulation.

after about 100 ps ( Figure 4.2). Only a slight increase in the Van der Waals energy is observed ( Figure 4.3). Finally, Figures 4.34 and 4.35 show the courses of the dispersion and repulsion energies, respectively. Both energies increase only slightly after 120 ps, when a more favourable electrostatic interaction has been reached.

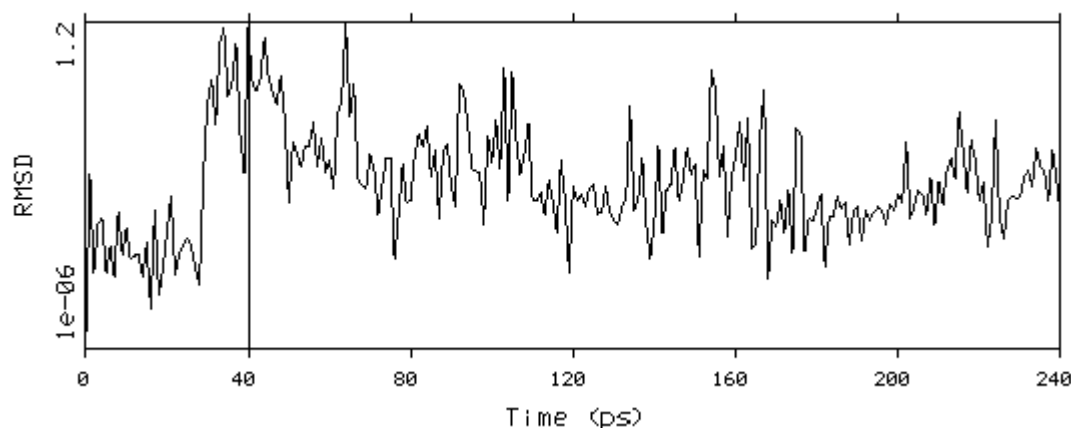


**Figure 4.4** Course of the dispersion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and oxacillin during 240 ps MD simulation.



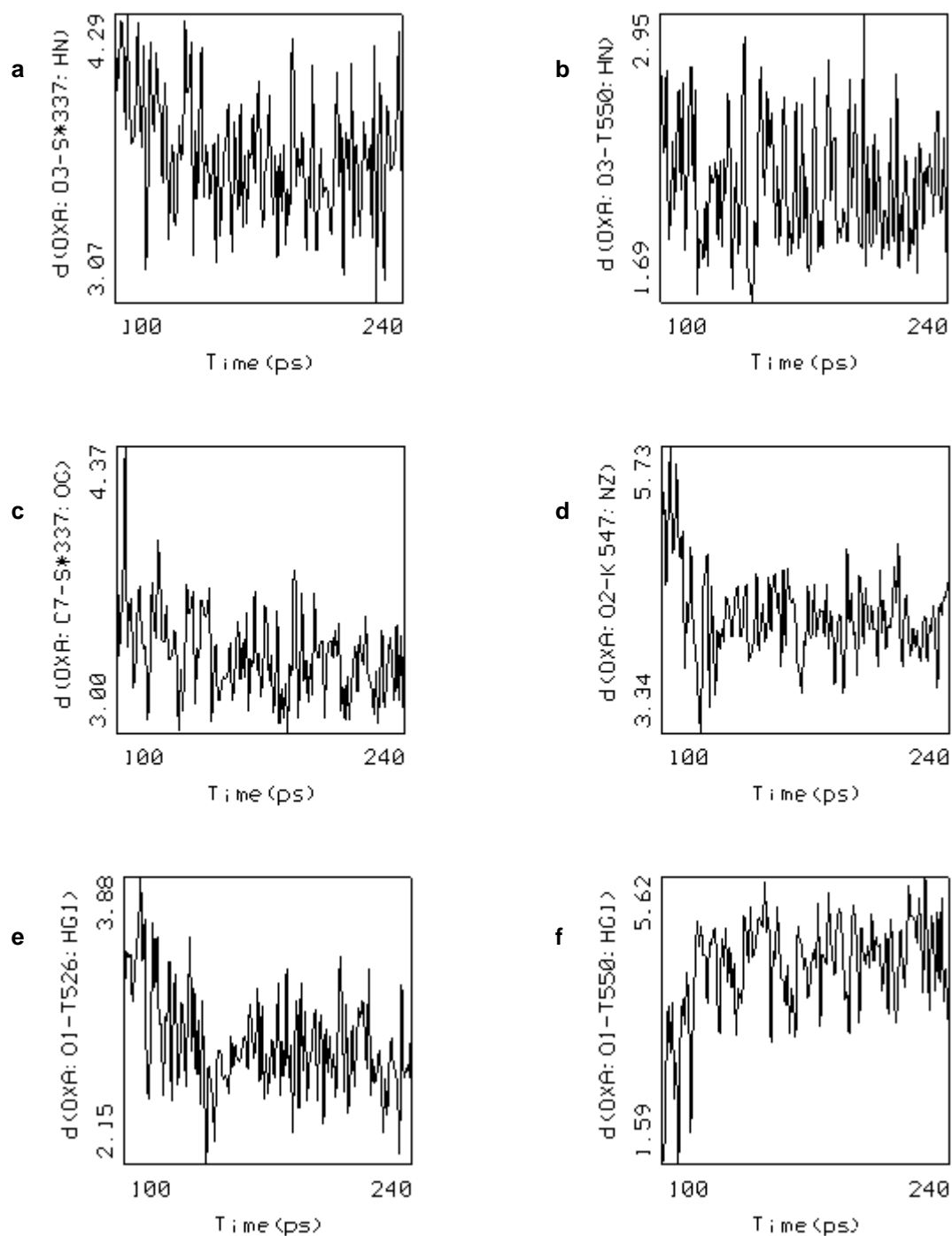
**Figure 4.5** Course of the repulsion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and oxacillin during 240 ps MD simulation.

The stability of the complex is further shown by the course of the RMS deviation of oxacillin during the simulation given in Figure 4.6.



**Figure 4.6** Course of the RMS deviation of oxacillin from the start configuration during the 240 ps simulation of the complex with *S. pneumoniae* PBP2x\* (strain R6).

Figure 4.7 shows the courses of some important intermolecular atomic distances during the last 140 ps.



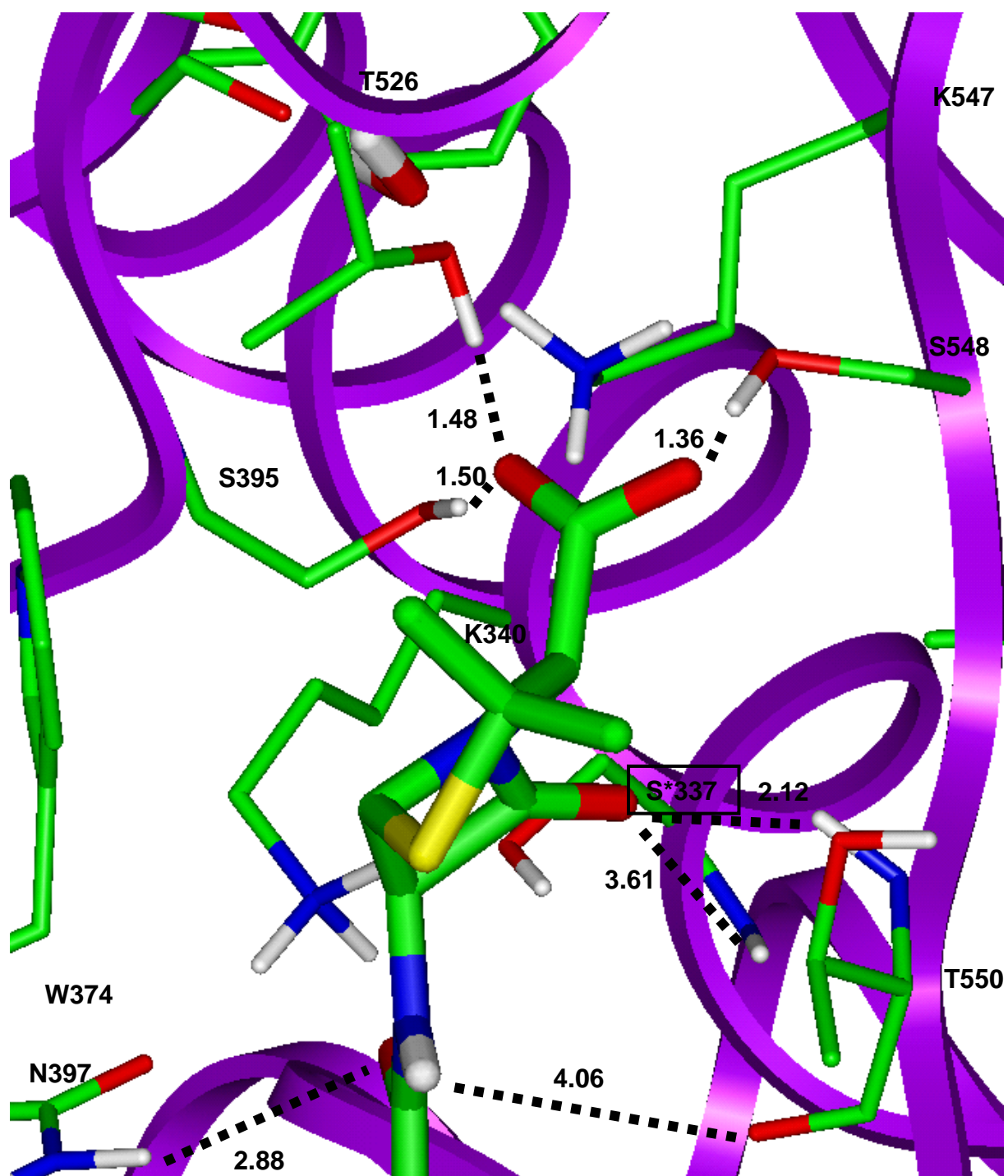
**Figure 4.7** Courses of various intermolecular atomic distances between PBP2x\* and oxacillin during the last 140 ps simulation of the complex. <sup>a</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of S\*337 <sup>b</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of T550 <sup>c</sup>Distance between the electrophilic lactam carbonyl carbon atom and the nucleophilic hydroxyl oxygen of S\*337 <sup>d</sup>Distance of the salt bridge between the charged nitrogen atom of K547 and the nearest carboxylate oxygen atom <sup>e</sup>Distance of the hydrogen bond between the hydroxyl hydrogen of T525 and the carboxylate group <sup>f</sup>Distance of the hydrogen bond between the hydroxyl hydrogen atom of T550 and the carboxylate group.

Again, Figure 4.8 presents the average structure of the complex during the last 140 ps simulation.



**Figure 4.8** Structure of PBP2x\* complexed with oxacillin, averaged over the last 140 ps of the MD simulation. The average structure of oxacillin is shown as atom type coloured sticks. The starting configuration is shown as thick black sticks, whereas the structures shown as small black sticks represent the conformation of oxacillin at 120 ps, 180 ps and 240 ps.

Finally, Figure 4.9 shows the Michaelis complex (average structure from 100 to 140 ps) zoomed in on the active site.

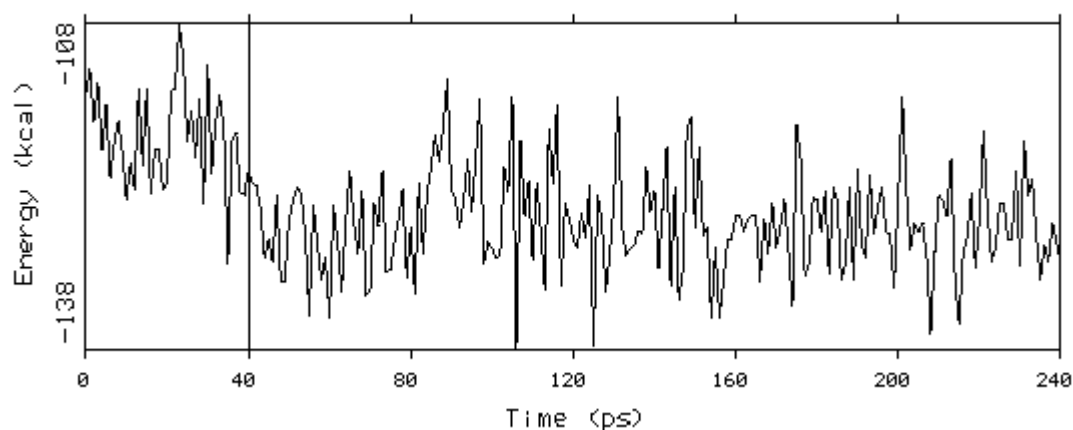


**Figure 4.9** Average structure of PBP2x\* complexed with oxacillin showing the hydrogen bonding interactions in the active site. Intermolecular distances shown represent average distances.

Like benzylpenicillin, the carboxylate group of oxacillin is highly stabilised by hydrogen bonding interactions with the side chains of S395, T526 and S548. The amide group is not in an optimal position to be maximally stabilised by hydrogen bonding interaction with the side chain of N397 and the backbone carbonyl oxygen of T550, either.

#### 4.5.2.5 Simulation of PBP2x\*<sub>R6</sub> complexed with cephalothin

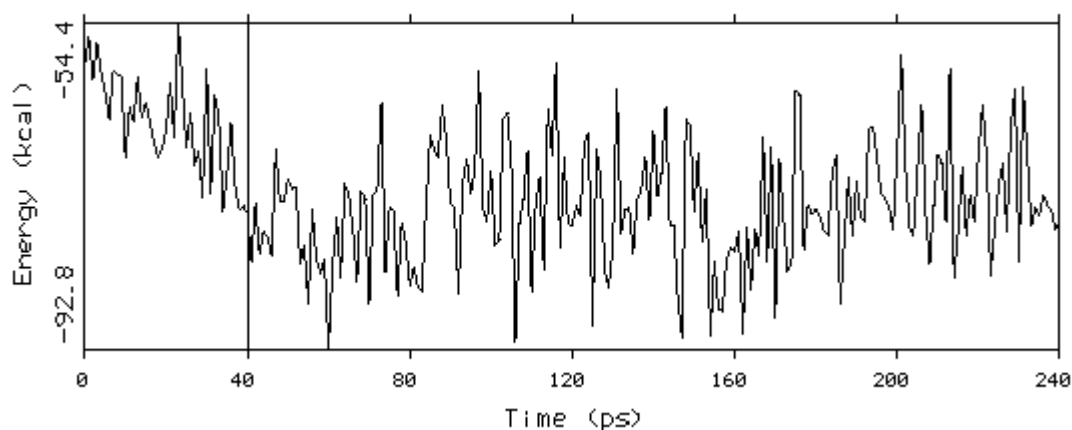
The same detailed analysis of the interaction energies of cephalothin complexed to PBP2x\* is given by the graphs shown in Figures 4.40 to 4.44. As follows from Figure 4.1 the interaction energy between the protein and the inhibitor remains constant throughout the simulation after 40 ps.



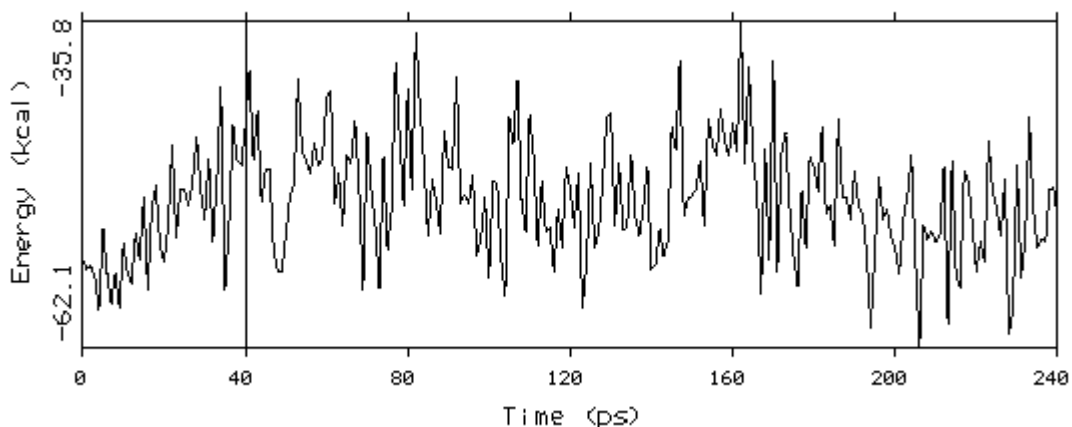
**Figure 4.1** Course of the total interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cephalothin during 240 ps MD simulation.

Also both the electrostatic and Van der Waals contributions to the interaction energy shown in Figures 4.41 and 4.42, respectively, remain constant.



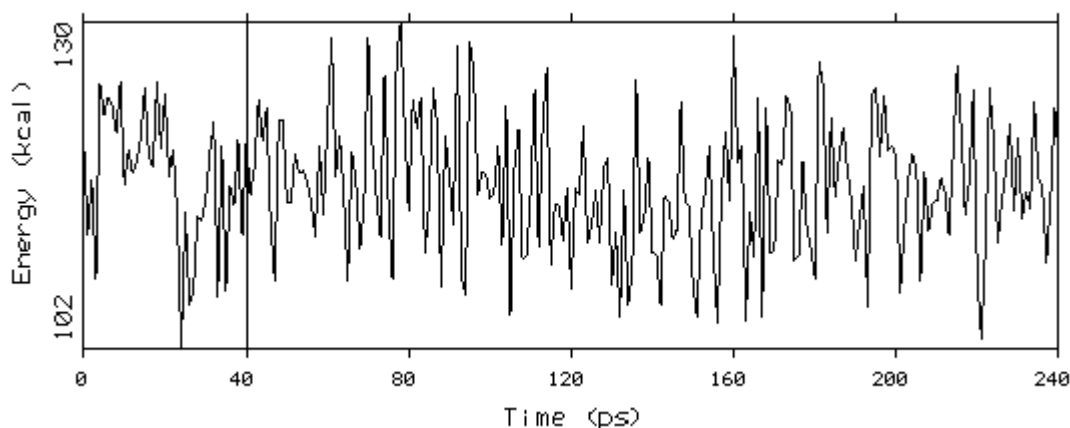


**Figure 4.2** Course of the Coulomb electrostatic contribution of the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cephalothin during 240 ps MD simulation.

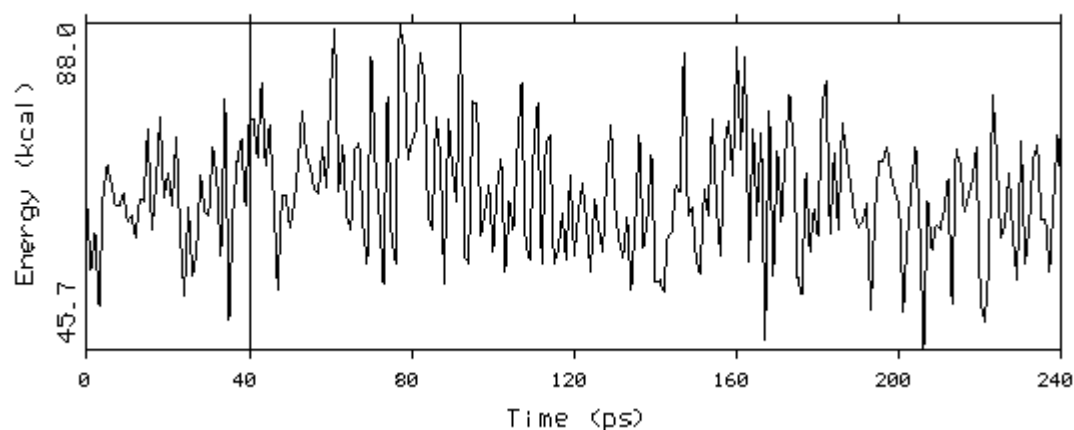


**Figure 4.3** Course of the Van der Waals contribution to the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cephalothin during 240 ps MD simulation.

Moreover, as shown in Figure 4.4 and 4.44, also both the dispersion and repulsion energies remain constant throughout the simulation.

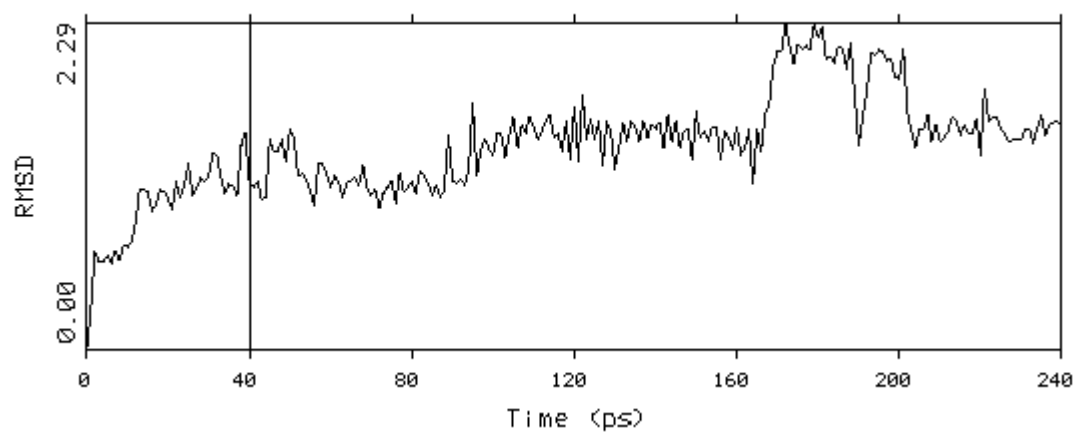


**Figure 4.4** Course of the dispersion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cephalothin during 240 ps MD simulation.

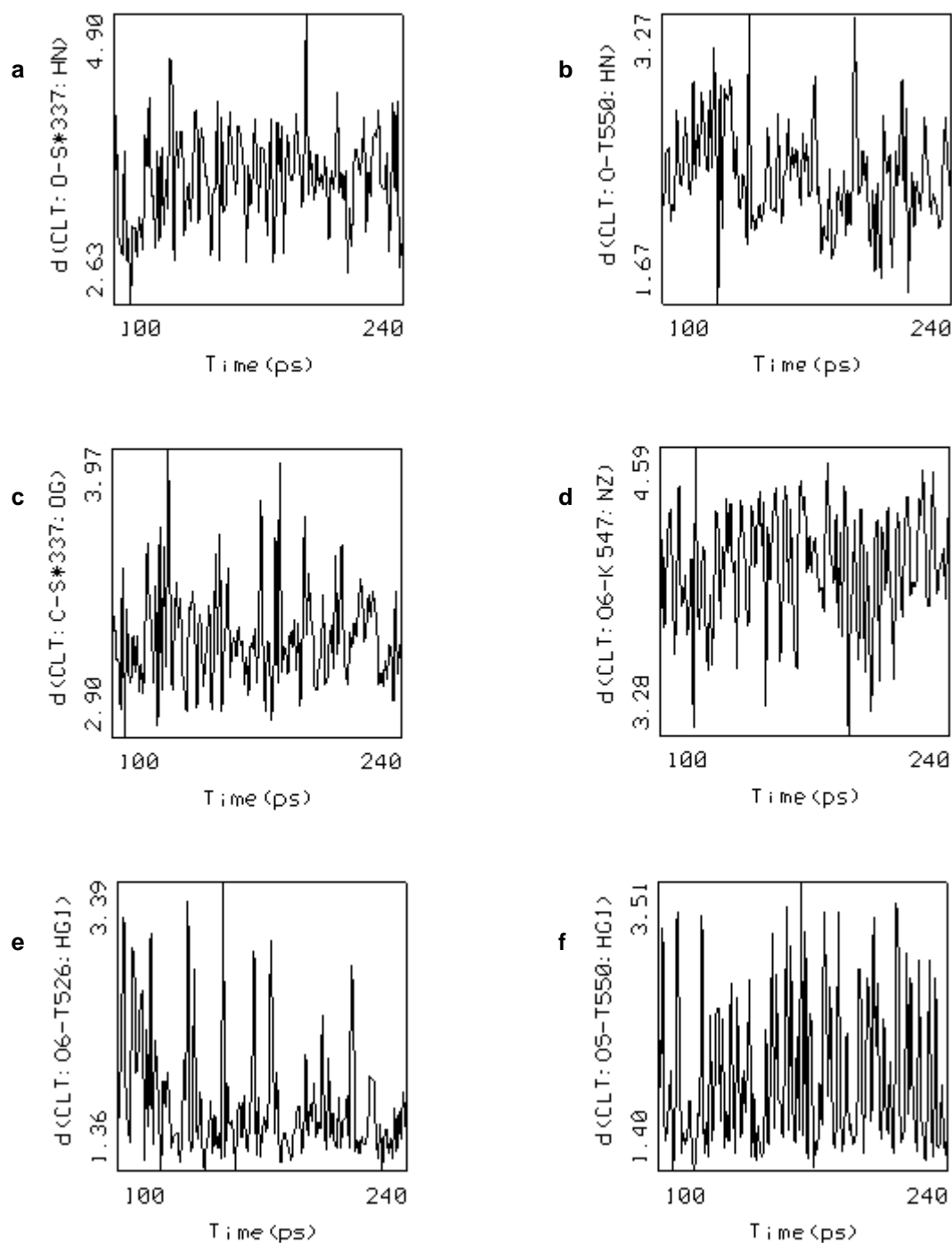


**Figure 4.5** Course of the Van der Waals repulsion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\* (R6 strain) and cephalothin during 240 ps MD simulation.

The stability of the complex is also reflected by the course of the RMS deviation of cephalothin during the simulation (Figure 4.6), although the ligand's position seems to shortly change in the second half of the simulation. As shown in Figure 4.8 this change is caused by movements of the hydrophobic thiophene substituent. The stability of the complex also follows from the courses of some important intermolecular atomic distances shown in Figure 4.7.



**Figure 4.6** Course of the RMS deviation of cephalothin from the start configuration during the 240 ps simulation of the complex with *S. pneumoniae* PBP2x\* (R6 strain).



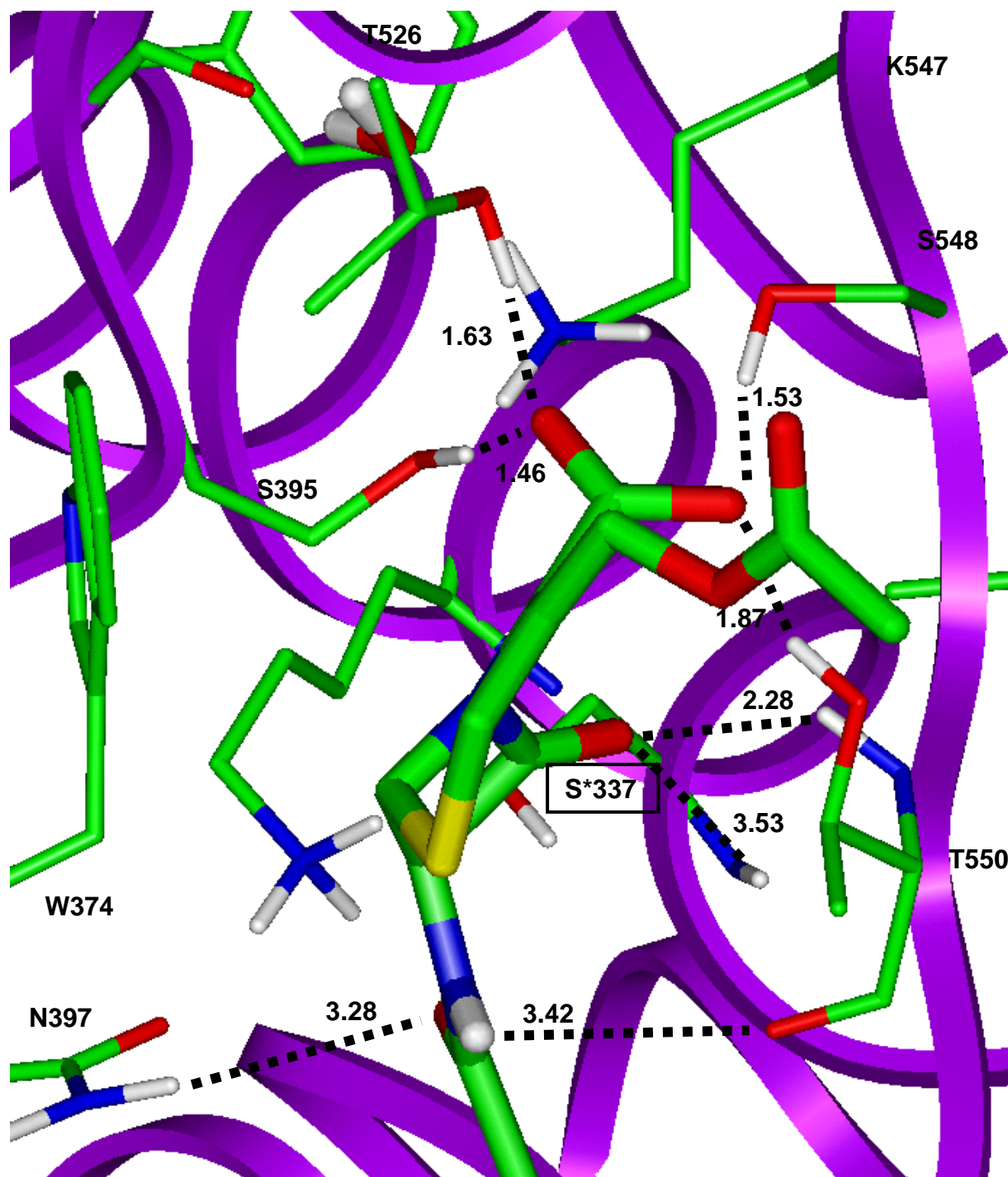
**Figure 4.7** Courses of various intermolecular atomic distances between PBP2x\* and cephalothin during the last 140 ps simulation of the complex. <sup>a</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of S\*337 <sup>b</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of T550 <sup>c</sup>Distance between the electrophilic lactam carbonyl carbon atom and the nucleophilic hydroxyl oxygen of S\*337 <sup>d</sup>Distance of the salt bridge between the charged nitrogen atom of K547 and the nearest carboxylate oxygen atom <sup>e</sup>Distance of the hydrogen bond between the hydroxyl hydrogen of T526 and the carboxylate group <sup>f</sup>Distance of the hydrogen bond between the hydroxyl hydrogen atom of T550 and the carboxylate group.

In order to only take the energetically more favourable second half of the simulation into account, Figure 4.8 presents the average structure of the complex during the last 140 ps simulation.



**Figure 4.8** Structure of PBP2x complexed with cephalothin, averaged over the last 140 ps of the MD simulation. The average structure of cephalothin is shown as atom type coloured sticks. The starting configuration is shown as thick black sticks, whereas the structures shown as small black sticks represent the conformation of cephalothin at 120 ps, 180 ps and 240 ps.

Figure 4.9 shows the Michaelis complex (average structure from 100 to 140 ps) zoomed in on the active site.



**Figure 4.9** Average structure of PBP2x\* complexed with cephalothin showing the hydrogen bonding interactions in the active site. Intermolecular distances shown represent average distances.

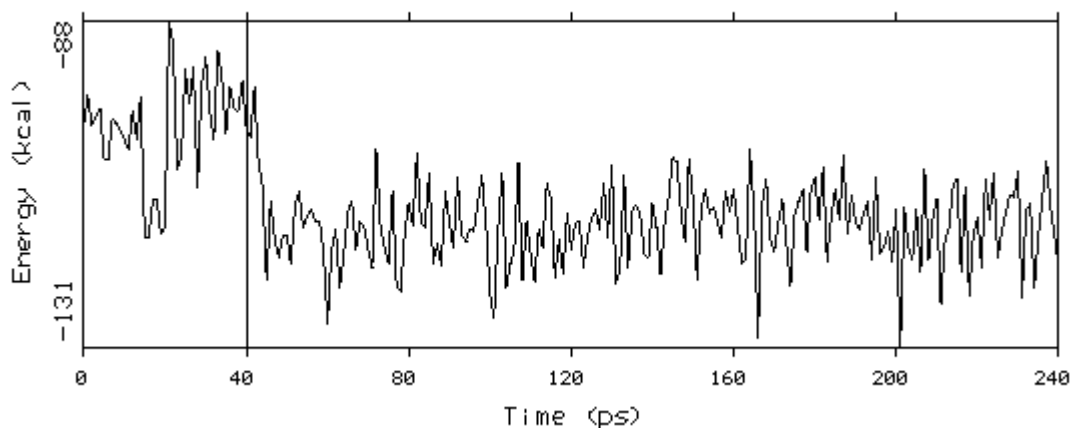
Like cefotaxime, the carboxylate group is stabilised by a strong hydrogen bonding interaction with the side chains of S395, T526 and S548. Due to the absence of stabilising hydrophilic interactions of the R1 substituent, though, the carboxylate group is not allowed to turn toward T550 as much as cefotaxime. As a result, a hydrogen bond with T526 is still possible. Contrasting to cefotaxime, however, the amide group is not optimally stabilised between N397 and T550.

### 4.5.3 Molecular dynamics simulations of PBP2x\*<sub>T526→S</sub> complexed with inhibitors

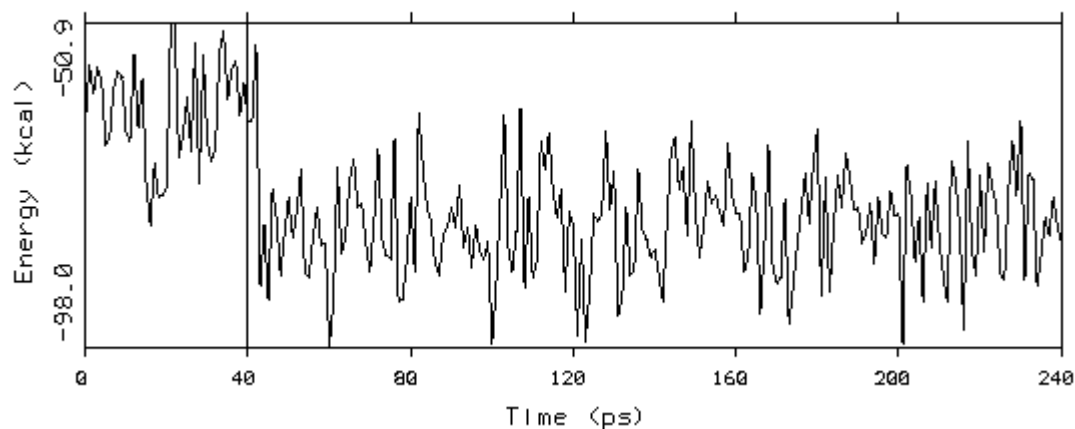
In order to investigate any influences of the T526→S mutation, T526 in both the starting conformations of the MD simulations performed of the R6 enzyme complexed with benzylpenicillin and cefotaxime was mutated into serine. Leaving both the inhibitor and the rest of the enzyme unaltered, the resulting complexes were again subjected to a 240 ps MD simulation.

#### 4.5.3.1 Simulation of PBP2x\*<sub>T526→S</sub> complexed with benzylpenicillin

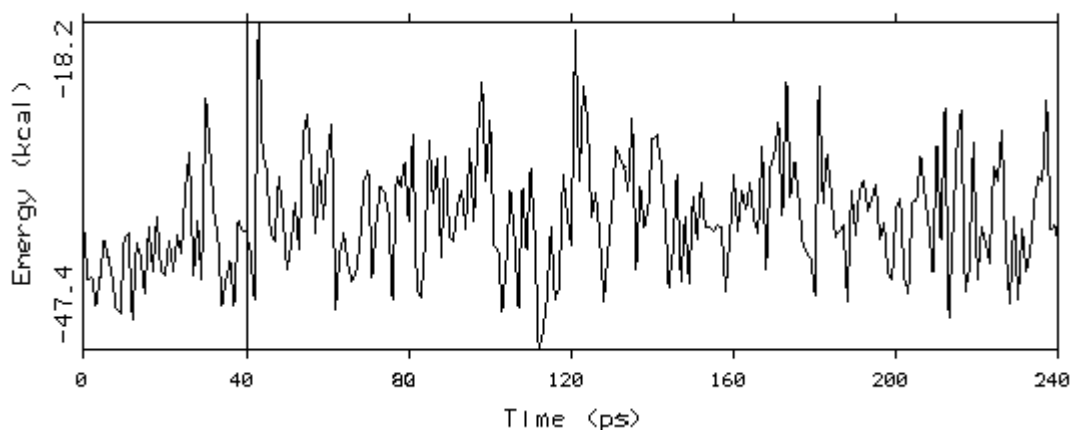
The trajectory graphs presented in Figure 4.1 to Figure 4.6 present a detailed analysis of the MD simulation of the PBP2x\*<sub>T526→S</sub> mutant protein complexed with benzylpenicillin. As follows from Figure 4.1 the interaction energy strongly decreases after 40 ps, to remain constant throughout the rest of the simulation. Also both the Coulomb and Van der Waals contributions to the interaction energy shown in Figure 4.2 and Figure 4.3, respectively, remain fairly constant. As shown in Figure 4.4 and Figure 4.5, also both the dispersion and repulsion energies remain fairly constant.



**Figure 4.1** Course of the total interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and benzylpenicillin during 240 ps MD simulation.



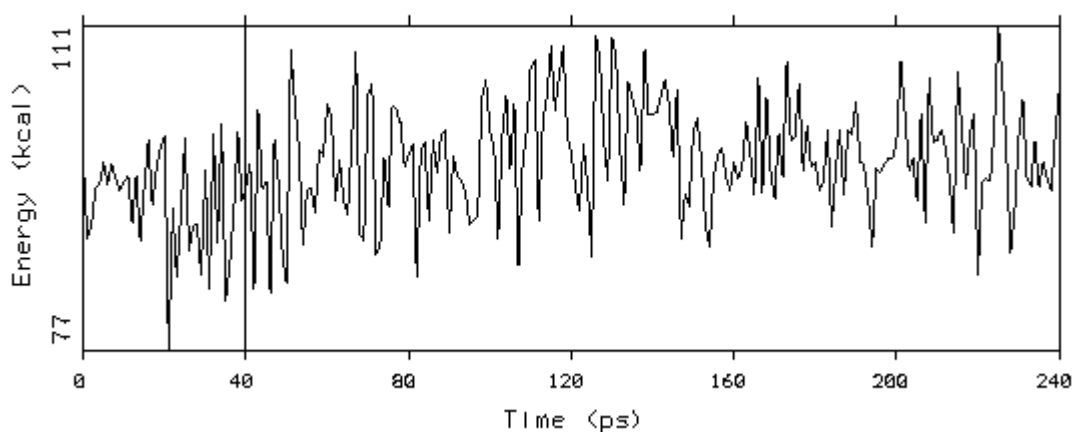
**Figure 4.2** Course of the Coulomb electrostatic contribution of the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and benzylpenicillin during 240 ps MD simulation.



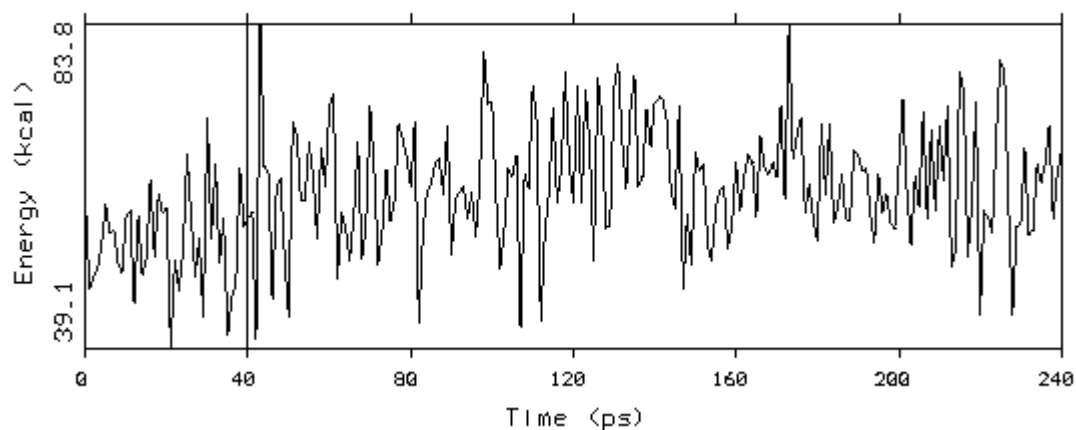
**Figure 4.3** Course of the Van der Waals contribution to the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and benzylpenicillin during 240 ps MD simulation.



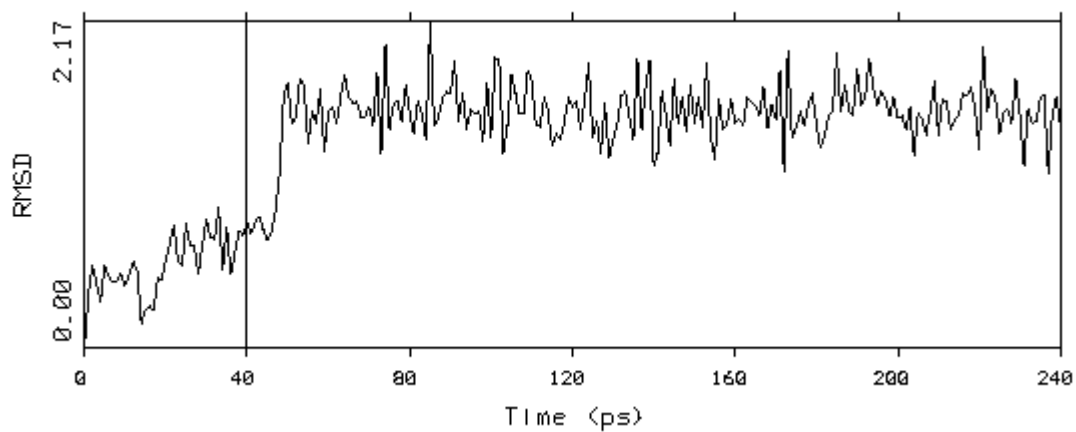
The stability of the complex is further shown by the course of the RMS deviation of benzylpenicillin during the simulation (Figure 4.6). Also the courses of some important intermolecular atomic distances (Figure 4.7) show that the complex is stable during the last 140 ps.



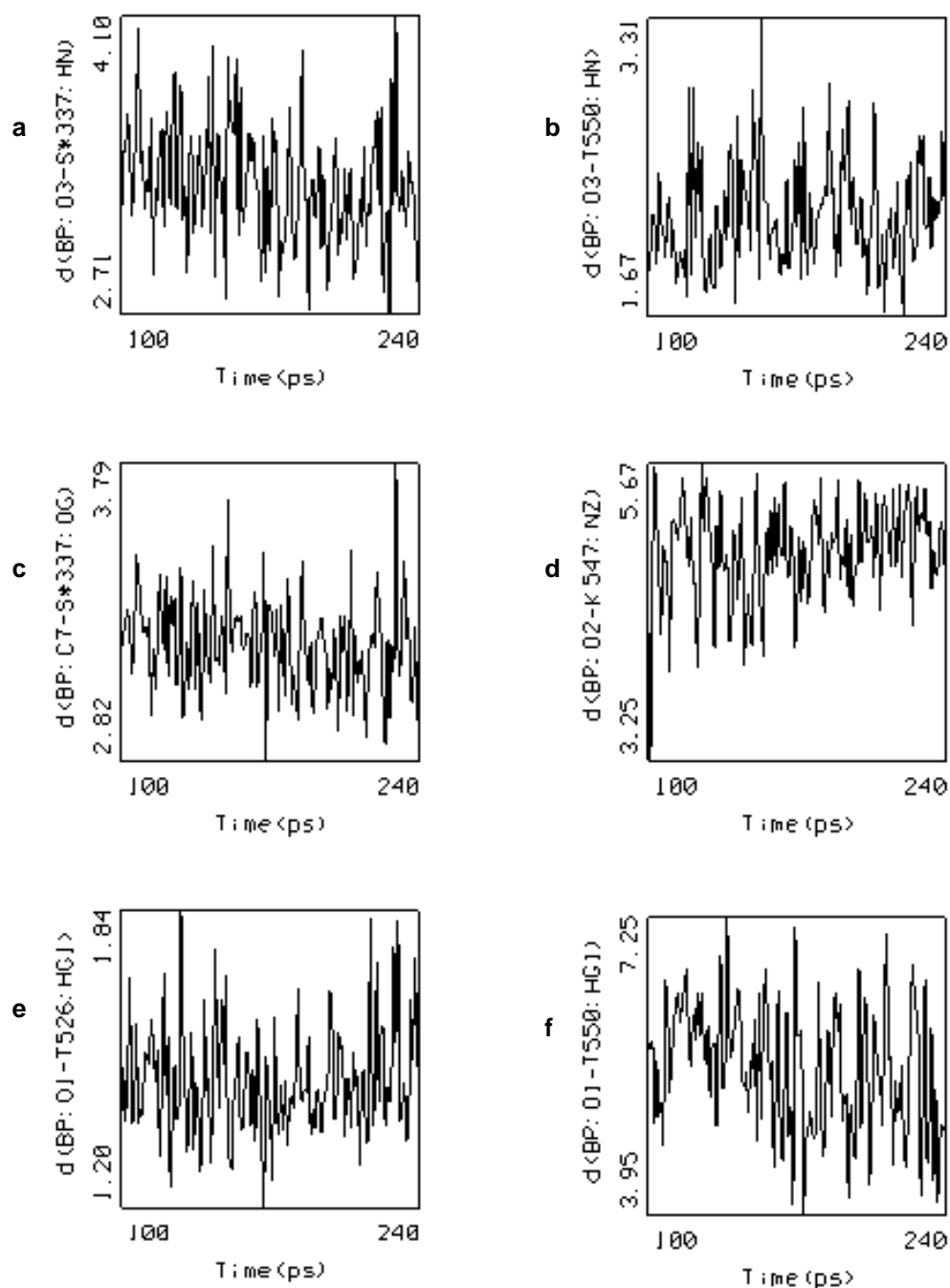
**Figure 4.4** Course of the Van der Waals dispersion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and benzylpenicillin during 240 ps MD simulation.



**Figure 4.5** Course of the Van der Waals repulsion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and benzylpenicillin during 240 ps MD simulation.



**Figure 4.6** Course of the RMS deviation of benzylpenicillin from the start configuration during the 240 ps simulation of the complex with *S. pneumoniae* PBP2x\*<sub>T526→S</sub>.



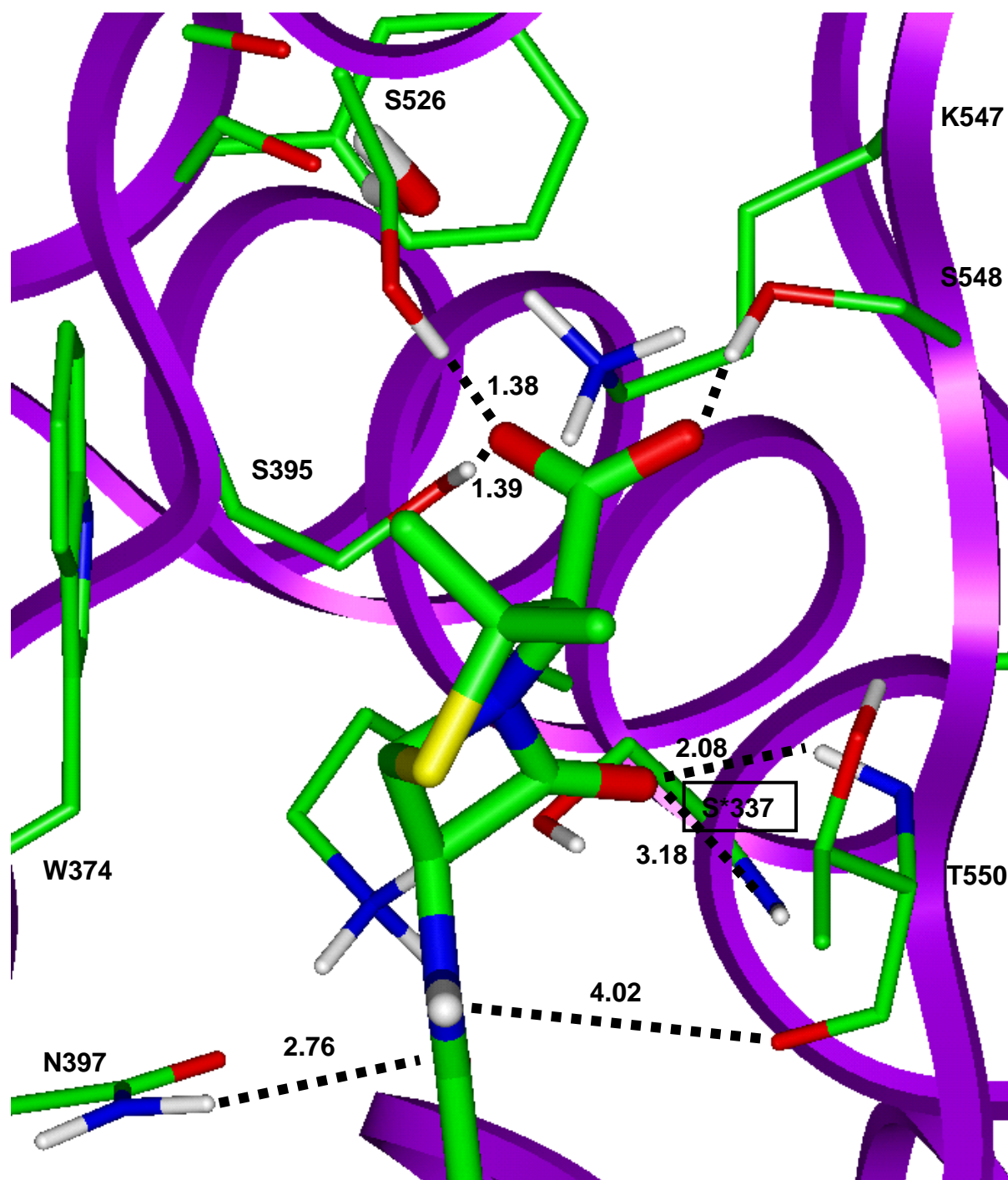
**Figure 4.7** Courses of various intermolecular atomic distances between  $PBP2x^*_{T526 \rightarrow S}$  and benzylpenicillin during the last 140 ps simulation of the complex. <sup>a</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of S\*337 <sup>b</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of T550 <sup>c</sup>Distance between the electrophilic lactam carbonyl carbon atom and the nucleophilic hydroxyl oxygen of S\*337 <sup>d</sup>Distance of the salt bridge between the charged nitrogen atom of K547 and the nearest carboxylate oxygen atom <sup>e</sup>Distance of the hydrogen bond between the hydroxyl hydrogen of T526 and the carboxylate group <sup>f</sup>Distance of the hydrogen bond between the hydroxyl hydrogen atom of T550 and the carboxylate group.

In order to be able to fully compare the results with the R6 enzyme, Figure 4.8 presents the average structure of the complex during the last 140 ps simulation.



**Figure 4.8** Structure of PBP2x\*<sub>T526→S</sub> complexed with benzylpenicillin, averaged over the last 140 ps of the MD simulation. The average structure of benzylpenicillin is shown as atom type coloured sticks. The starting configuration is shown as thick black sticks, whereas the structures shown as small black sticks represent the conformation of benzylpenicillin at 120 ps, 180 ps and 240 ps.

Figure 4.9 shows the Michaelis complex (average structure from 100 to 140 ps) zoomed in on the active site.



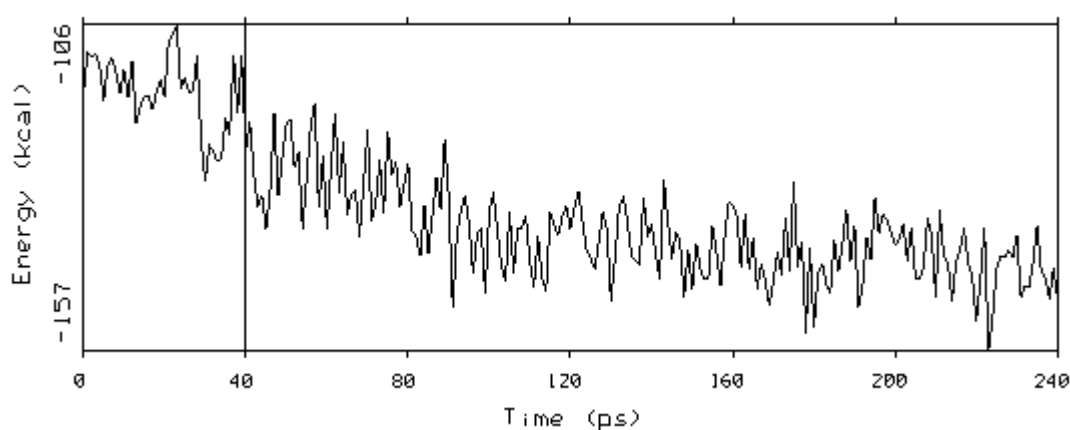
**Figure 4.9** Average structure of  $PBP2x_{T526 \rightarrow S}$  complexed with benzylpenicillin showing the hydrogen bonding interactions in the active site. Intermolecular distances shown represent average distances.

Like in the R6 enzyme, the carboxylate group is stabilised by the same three hydrogen bonds. However, since rotation of S526 is now unhindered, the  $\gamma$ -OH group is now able to bridge from a geometrically different position. Since this seems to

cause the inhibitor to be tilted in the active site compared to the R6 complex, optimal stabilisation of the amide group is not possible anymore.

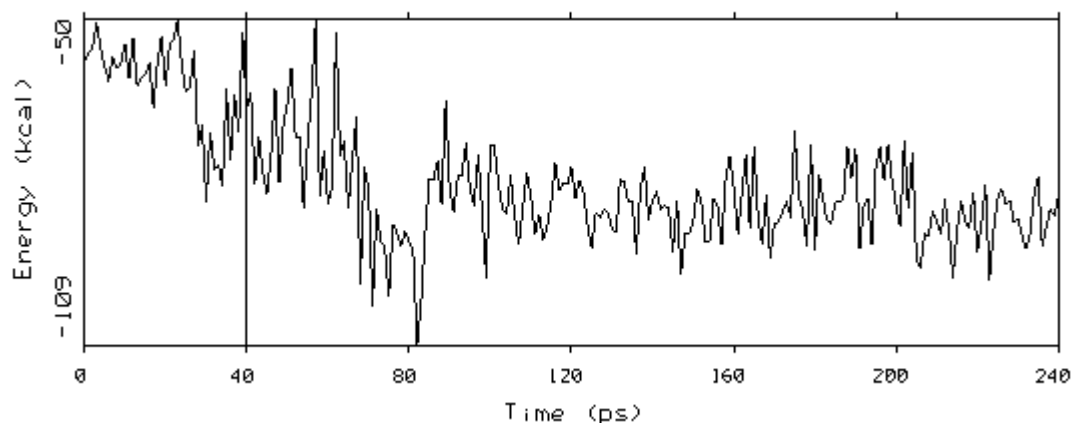
#### 4.5.3.2 Simulation of PBP2x\*<sub>T526→S</sub> complexed with cefotaxime

The same detailed analysis of the interaction energies of cefotaxime complexed to PBP2x<sub>T526→S</sub> is given by the graphs shown in Figure 4.1 to Figure 4.6. As follows from Figure 4.1 the interaction energy between the mutant protein and the inhibitor remains fairly constant throughout the simulation after 40 ps.

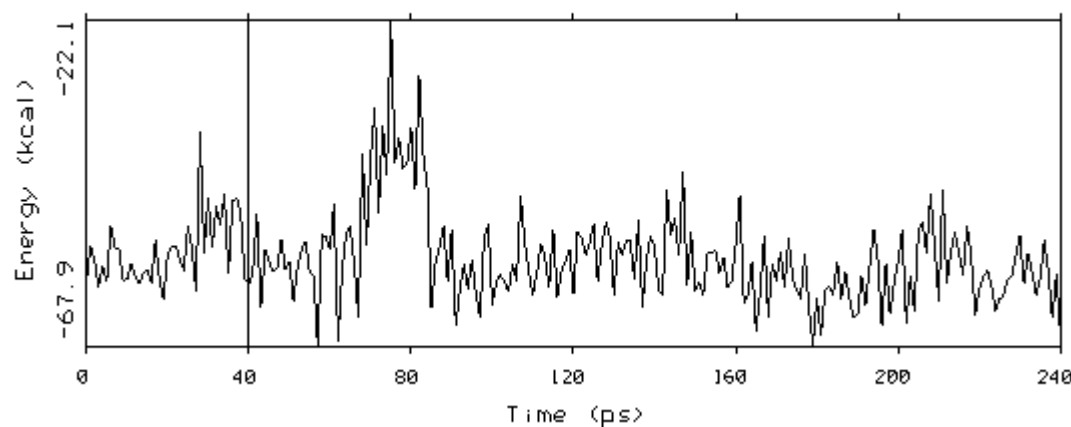


**Figure 4.1** Course of the total interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and cefotaxime during 240 ps MD simulation.

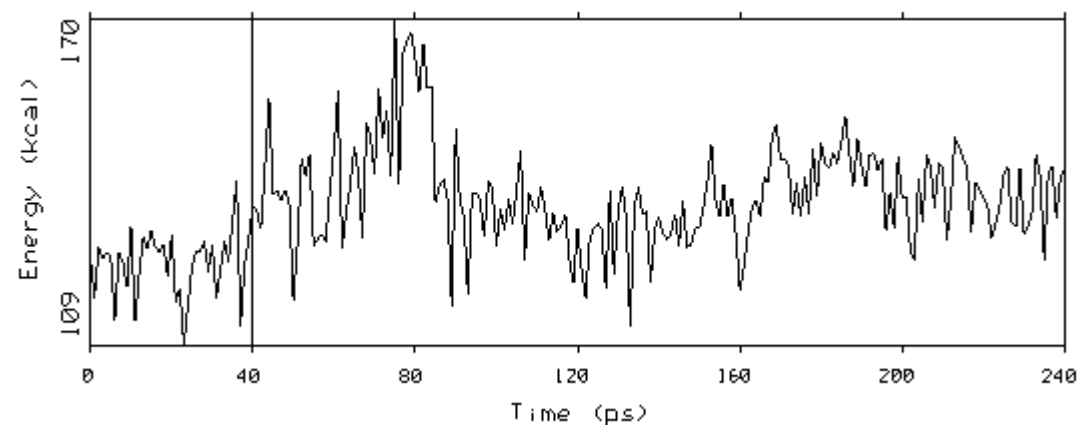
Also both the electrostatic and Van der Waals contributions to the interaction energy shown in Figure 4.2 and Figure 4.3, respectively, remain constant after the initial 40 ps. Moreover, as shown in Figure 4.4 and Figure 4.5, also both the dispersion and repulsion contributions to the Van der Waals energy remain fairly constant.



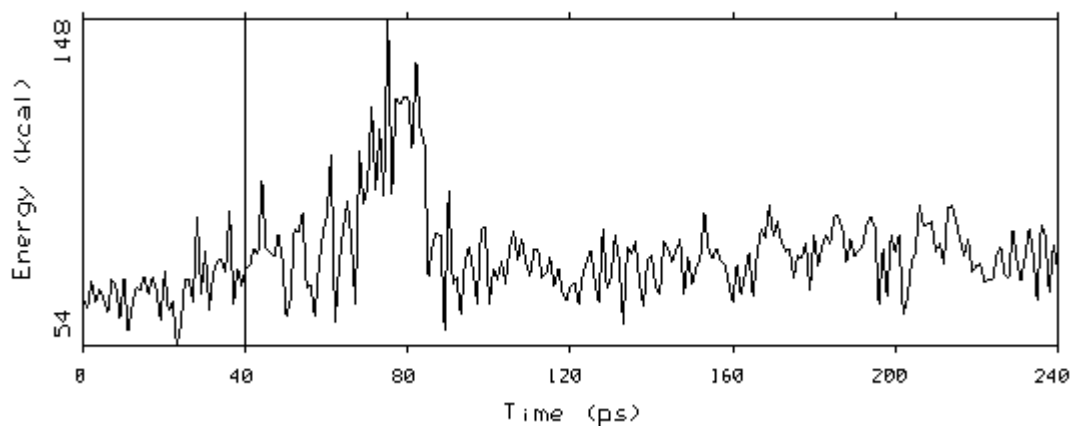
**Figure 4.2** Course of the Coulomb electrostatic contribution of the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and cefotaxime during 240 ps MD simulation.



**Figure 4.3** Course of the Van der Waals contribution to the interaction energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and cefotaxime during 240 ps MD simulation.

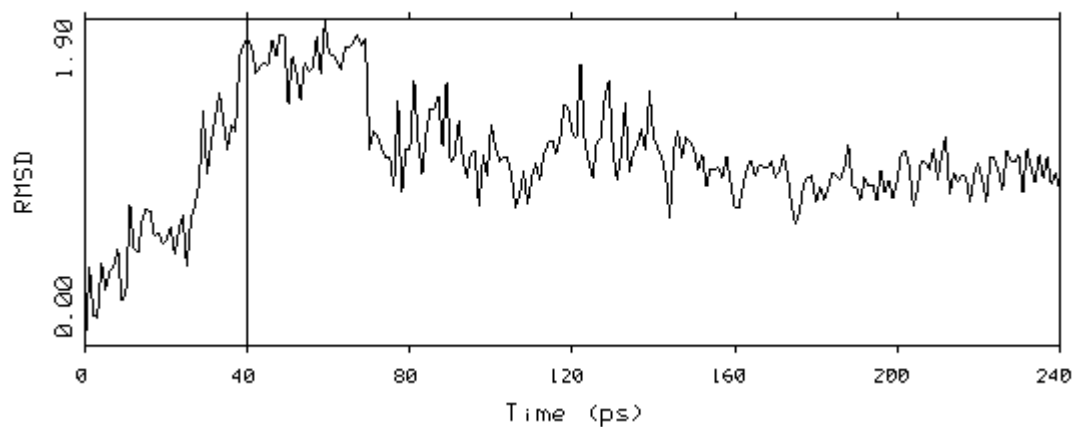


**Figure 4.4** Course of the Van der Waals dispersion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and cefotaxime during 240 ps MD simulation.



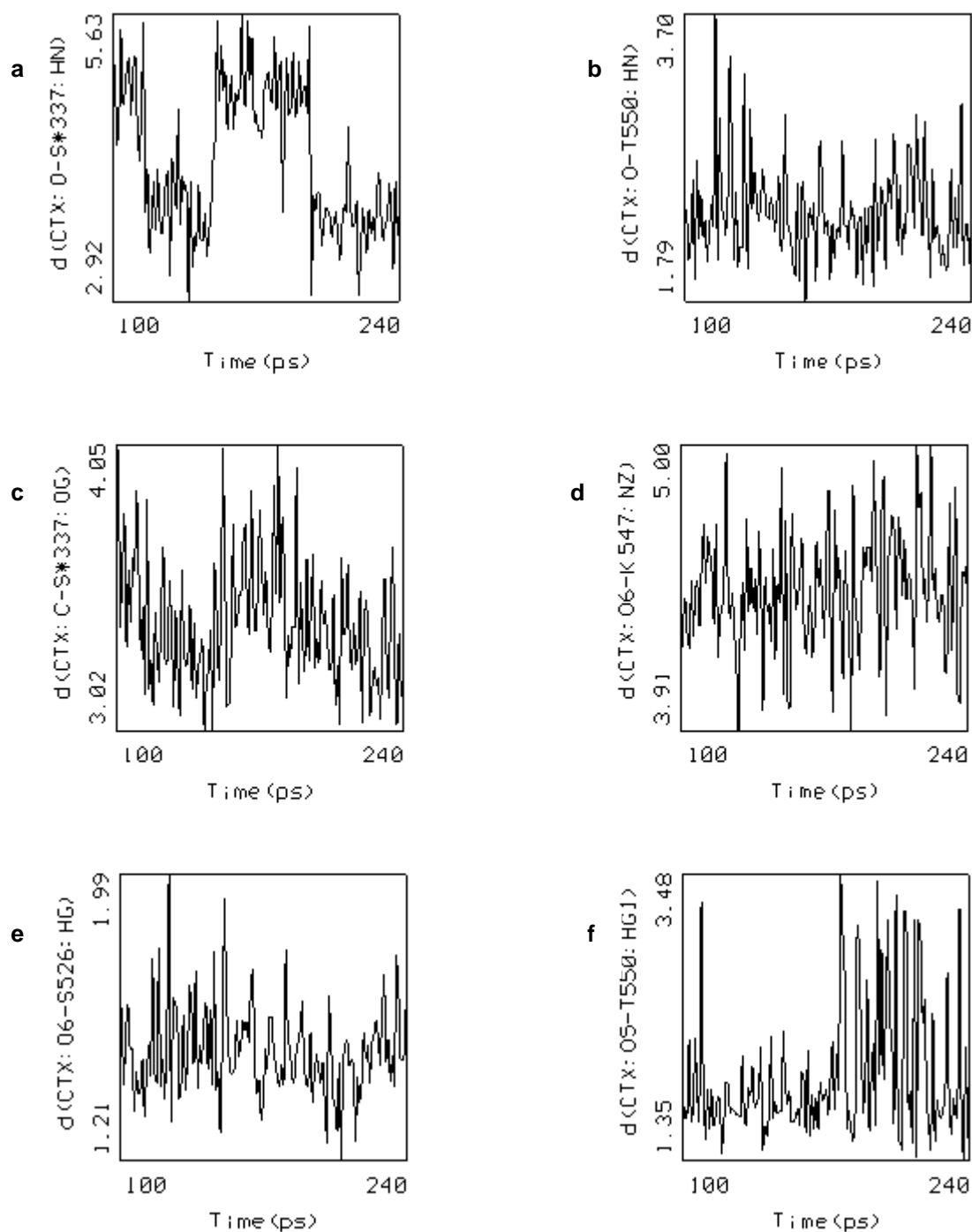
**Figure 4.5** Course of the Van der Waals repulsion energy between the transpeptidase domain of *S. pneumoniae* PBP2x\*<sub>T526→S</sub> and cefotaxime during 240 ps MD simulation.

The stability of the complex is further shown by the course of the RMS deviation of cefotaxime during the simulation ( Figure 4.6).



**Figure 4.6** Course of the RMS deviation of cefotaxime from the start configuration during the 240 ps simulation of the complex with *S. pneumoniae* PBP2x\*<sub>T526→S</sub>.





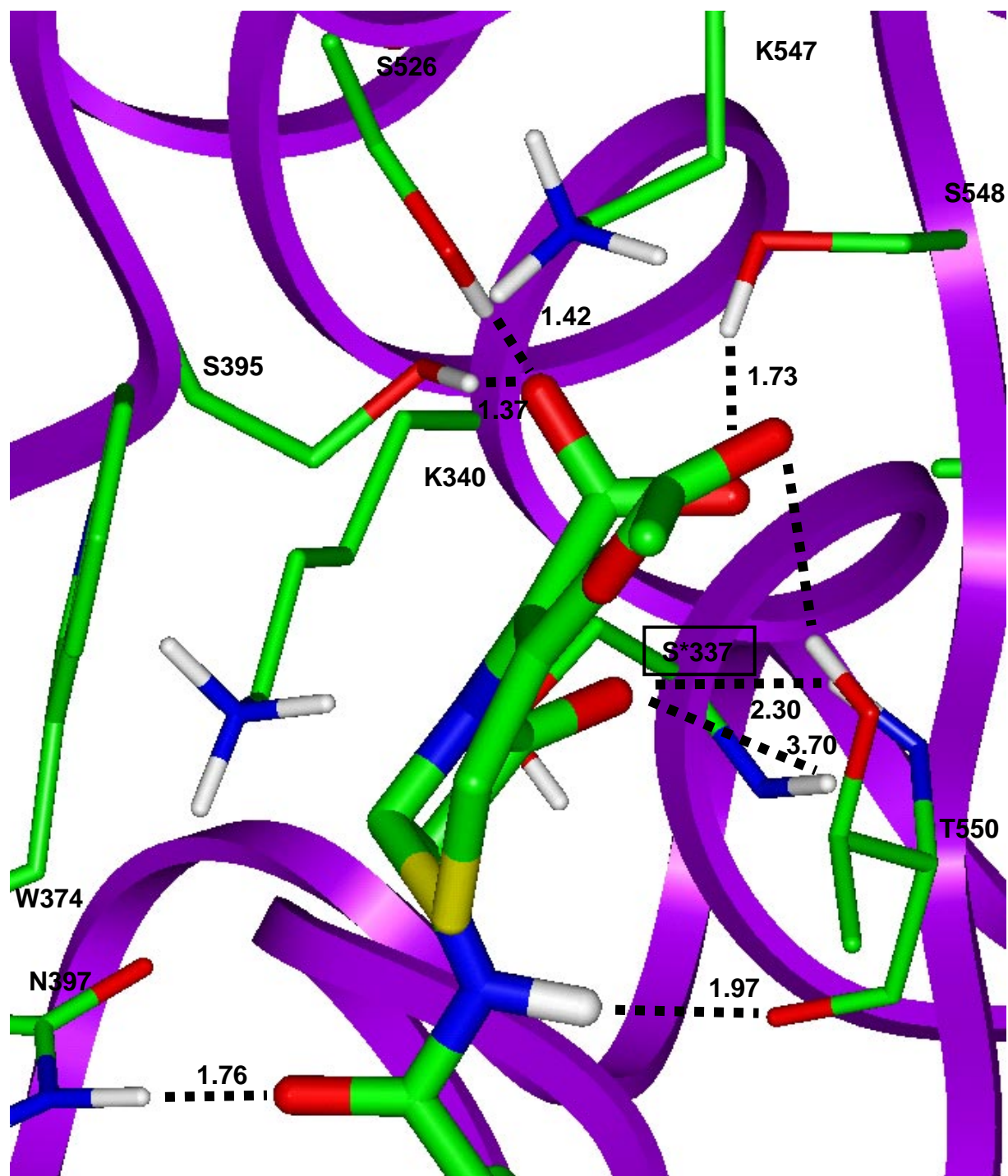
**Figure 4.7** Courses of various intermolecular atomic distances between  $PBP2x^*_{T526-S}$  and cefotaxime during the last 140 ps simulation of the complex. <sup>a</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of S\*337 <sup>b</sup>Distance of the hydrogen bond between the lactam carbonyl oxygen and the backbone NH of T550 <sup>c</sup>Distance between the electrophilic lactam carbonyl carbon atom and the nucleophilic hydroxyl oxygen of S\*337 <sup>d</sup>Distance of the salt bridge between the charged nitrogen atom of K547 and the nearest carboxylate oxygen atom <sup>e</sup>Distance of the hydrogen bond between the hydroxyl hydrogen of T526 and the carboxylate group <sup>f</sup>Distance of the hydrogen bond between the hydroxyl hydrogen atom of T550 and the carboxylate group.

In order to only take the energetically more favourable second half of the simulation into account, Figure 4.8 presents the average structure of the complex during the last 140 ps simulation.



**Figure 4.8** Structure of PBP2x\*<sub>T526→S</sub> complexed with cefotaxime, averaged over the last 140 ps of the MD simulation. The average structure of benzylpenicillin is shown as atom type coloured sticks. The starting configuration is shown as thick black sticks, whereas the structures shown as small black sticks represent the conformation of cefotaxime at 120 ps, 180 ps and 240 ps.

Figure 4.9 shows the Michaelis complex (average structure from 100 to 140 ps) zoomed in on the active site.



**Figure 4.9** Average structure of  $PBP2x^*_{T526 \rightarrow S}$  complexed with cefotaxime showing the hydrogen bonding interactions in the active site. Intermolecular distances shown represent average distances.

In the mutant protein the carboxylate group is stabilised by the same hydrogen bonds as benzylpenicillin. Contrasting to the latter inhibitor, however, the amide group is now optimally stabilised between N397 and T550.

#### 4.5.4 Summary of the various simulations

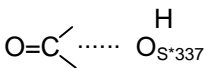
Although the absolute energy values derived from force field simulations do not have any physical meaning, as explained in section 3.1.3.3 the dispersion energy values are especially suited for quantifying the polarisability. Thus, these values should be able to give clues as for the ease with which acylation step proceeds. Table 4. 9 lists the total, coulomb, Van der Waals, dispersion and repulsion energy values of the various MD simulations of the complexes run with both the R6 and T526→S enzymes, averaged over the last 140 ps.

**Table 4. 9** *Summary of the various molecular dynamics simulations of the complexes run with PBP2x\*<sub>R6</sub> (section 4.5.2) and PBP2x\*<sub>T526→S</sub> (section 4.5.3). Values are in kcal/mol.*

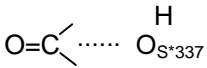
		Total energy	Coulomb energy	Van der Waals energy	Dispersion energy	Repulsion energy
R6	bp	-113.7	-75.9	-37.7	97.7	60.0
	ctx	-126.2	-76.6	-49.6	118.3	68.7
	oxa	-120.1	-73.2	-46.9	109.9	63.0
	clt	-128.6	-79.9	-48.6	112.7	64.7
T526→S	bp	-115.2	-79.8	-35.4	98.0	92.7
	ctx	-128.1	-71.2	-56.9	135.8	78.9

Likewise, some important geometrical values of the Michaelis complexes of the various antibiotics derived by the dynamics simulations are summarised in Table 4. 10.

**Table 4. 10** Summary of the most important intermolecular atomic distances of the Michaelis complexes as derived from the molecular dynamics simulations of the complexes. The values presented represent the average values during the last 140 ps of the simulation.

	Oxyanion hole		Carboxylate		Reactive atom pair
	$C=O \cdots HN_{S^{*337}}$	$C=O \cdots HN_{T550}$	$COO^{-} \cdots HO_{T526}$	$COO^{-} \cdots HO_{T550}$	
bp	3.57	2.24	1.48	---	3.30
ctx	3.09	2.69	---	1.51	3.16
oxa	3.61	2.12	1.50	---	3.35
clt	3.53	2.28	1.63	1.87	3.22

	Oxyanion hole		Carboxylate		Reactive atom pair
	$C=O \cdots HN_{S^{*337}}$	$C=O \cdots HN_{T550}$	$COO^{-} \cdots HO_{S526}$	$COO^{-} \cdots HO_{T550}$	
bp	3.18	2.08	1.38	3.56	3.12
ctx	3.70	2.30	1.42	1.79	3.20

Finally, Table 4. 11 summarises in which complexes the amide group is stabilised between the side chain of N397 and the backbone carbonyl group of T550, and in which not.

**Table 4. 11** Summary in which complexes the amide group is stabilised between N397 and T550 and in which not (bp = benzylpenicillin, ctx = cefotaxime, oxa = oxacillin, clt = cephalothin).\*

---

	bp	ctx	oxa	clt
PBP2 <sub>X<sub>R6</sub></sub>	no	yes	no	no
PBP2 <sub>X<sub>T526→S</sub></sub>	no	yes	ND	ND

---

\* ND = not determined

## **5 Discussion**

---





## 5.1 Introductory aspects

The aim of this thesis was to shed light on the effects of some natural mutations in the active site region of *S. pneumoniae* PBP2x on the binding behaviour of  $\beta$ -lactam antibiotics. It should be stressed, however, that it is impossible to study the full process of acylation (and deacylation) with standard force field methods, because these methods fail to describe chemical processes like bond breaking and formation. However, since it is almost impossible to obtain detailed experimental information as to the process of forming a covalent acylenzyme complex either, the only way to gain some insight into structural properties governing the acylation process with force field simulations is to perform modelling studies of the regions surrounding the active site of the protein complexed with ligands. A prerequisite for performing such studies, though, is the availability of a reliable model of the native enzyme. Only then it should be possible to give a clue as to whether some amino acid mutations in the active site region result in a decreased affinity for a particular antibiotic or not.

In the work presented here, first a full three-dimensional model of PBP2x was developed using three-dimensional information of the structurally related  $\beta$ -lactamases, after which different antibiotics were docked into the active site and molecular dynamics simulations were run of the complexes formed.

Shortly after completion of the MD simulations performed of the modelled transpeptidase domain complexed with the four inhibitors new crystal structures were deposited in the PDB[191]. Although the corresponding article was not yet available, it was possible to compare the protein structure with both of the newly deposited crystal structures with PDB codes 1QME (2.4 Å; complex with sulfate) and 1QMF(2.7 Å; covalent complex with cefuroxime). Comparison of the model generated here and the crystal structures showed that particularly the active site region was modelled correctly.

## 5.2 The procedure for generating full structure from C $^{\alpha}$ information

Section 3.3.6 describes some methods for generating full protein coordinates starting from C $^{\alpha}$  atom information. Each of these well-accepted generally applicable methods generate full protein structures, which exactly match the C $^{\alpha}$  atom positions. However, since the optimisation of protein structure is a highly complicated multidimensional task, problems are likely to arise, if these experimental data correspond to poorly resolved structures. Like most methods for generating full protein structures from C $^{\alpha}$  atom information, in trying to avoid such problems arising from badly modelled starting structures, the approach developed here concentrates on optimisation of the the backbone hydrogen bonding network. Indeed, it is well-recognised [167] that the backbone geometry is mainly governed by hydrogen-bonding interactions, while a crude protein structure might be determined by forces optimising the hydrophobic packing interaction.

For these reasons the procedure presented here starts with local trace alignments of secondary structure elements and some parts of loops. These structurally aligned segments are incorporated into an initial model of the protein (which can be generated with any protein modelling package) by folding the corresponding regions. Subsequently, all amino acids, except for the glycine and proline residues, are mutated into alanine residues, after which the structure is subjected to a simulated annealing molecular dynamics simulation. This procedure was chosen, since glycine shows considerable differences in geometrical backbone properties compared to other amino acids, and proline is not able to act as a hydrogen bond donator. This stage of the procedure may be seen as an extension of the method described by Claessens et al. [153], who built a polyalanine peptide by searching a crystallographic database for optimally matching segments, and used the resulting protein as a starting structure for MD simulations.

Because the method starts with a local trace alignment, there are also similarities with the fragment-matching methods developed by Jones and Thirup [149], Reid and Thornton [150], and Holm and Sander [151]. Likewise, in order to match a fragment of varying length starting at a random C $^{\alpha}$ -position Levitt [154] searched a database of highly refined structures. By repeating this procedure several times he generated a

set of different protein structures matching the trace, which were averaged afterwards. However [160a, 150, 151, 162, 192, 193], problems may arise at junctions by overlapping one or two peptide groups. Such difficulties can be avoided, by simply folding the corresponding regions of a previously generated initial model according to the aligned regions.

In addition, by refolding the *whole* protein according to the geometry of the trace of the crystal structure, the *angles* and *dihedral angles* of the trace of the crystal structure could be recovered completely. The fact that readjusting these angles, the C-terminal domain and the transpeptidase domain itself, had resulted in slightly displaced positions with respect to the crystal structure could be addressed by modification of the distance between two consecutive C $\alpha$ -atoms along the virtual chain bond vector. This aspect of the procedure could be of general value in homology modelling, for example in the procedure of connecting secondary structure elements by the insertion of loops.

Also the relaxation procedure developed here has been proven to suit very well. With the purpose of generating full protein coordinates starting from C $\alpha$  atom information and making use of homology modelling techniques, it is apt to conserve their coordinates throughout the relaxation procedure. For this reason it is desirable to slowly relax the 3D homology derived regions in a positionally constrained framework, while simultaneously allowing for enough flexibility in both kinds of regions. Therefore, different force constants are needed. Whereas a single tethering force is divided between *all* the atoms in question, application of a restraining force allows one to selectively restrain atoms with different user-defined forces.

Since a number of arbitrary decisions have to be made during the course of the method (as is the case with most automated protein prediction methods [194]), though, it is hardly possible to judge the method quantitatively. The quality of the final model, therefore, strongly depends on the intuition and skill of the modeller. The only way to judge the quality of the method described here is comparison of the final model with a better solved X-ray structure. Penicillin binding protein 2x (PBP2x) of *Streptococcus pneumonia* served as an excellent example for this purpose.

## 5.3 Generation of the transpeptidase domain of *S. pneumoniae* PBP2x\*

### 5.3.1 Local structural trace alignment

Because the protein sequence and the coordinates of the C $^{\alpha}$  atoms were the only available experimental data of PBP2x, it was decided to perform a local structural trace alignment of the secondary structure elements with the structurally most related proteins with known 3D structure, i.e. the five  $\beta$ -lactamases. Although the functions of PBPs and  $\beta$ -lactamases are not the same, both interact with  $\beta$ -lactam antibiotics, and at least some measure of *functional* homology was incorporated in the alignment.

### 5.3.2 Relaxation of the backbone

After completion of the inevitable folding procedure discussed in section 0, the resulting hydrogen bonding network stabilising the five-stranded  $\beta$ -sheet was not optimal. In an attempt to optimise this network, molecular dynamics simulations of a strongly simplified poly-Gly-Ala-Pro protein were performed at 1000 K. Thus, peptide groups as a whole were allowed to rotate, without being hindered by steric clashes of side chains. Again, the use of restraining forces seemed to be more sophisticated to maintain the overall protein conformation, because the application of these forces enabled restriction of the C $^{\alpha}$  atoms to their crystal structure positions very effectively.

The idea of performing molecular dynamics studies on proteins by concentrating on backbone conformation is not new. For example, Levy *et al.* [195] introduced chain-restoring interactions in their quasi-harmonic model for studying the protein backbone fluctuations without the need of having to perform heavy calculations. Higo and Umeyama [196] even extended this backbone model with C $^{\alpha}$ -atom-packing interactions in order to study protein low-frequency vibrational modes. Likewise, in the backbone dynamics method developed here, it could also be nicely shown, that some parts of the protein were more flexible than other regions.

### 5.3.3 Relaxation of the whole transpeptidase domain

The data presented in Tables 4.4 and 4.5 show that the relaxation protocol used here served very well. The high percentage of amino acids in the most favoured regions of 79 - 82 % is maintained throughout all restrained dynamics, whereas the internal energy decreases steadily. Moreover, taking into account that the RMS deviation of the C<sup>α</sup>-atoms in the transpeptidase domain of the *optimised* structures reflect the error of the 351 C<sup>α</sup>-atoms of the transpeptidase domain of the model with respect to those of a crystal structure solved at 3.5 Å resolution, it may be concluded that the coordinates of the trace atoms are maintained very well.

Nevertheless, the RMS deviation of the corresponding 351 C<sup>α</sup>-atoms from those contained in the 3.5 Å resolution crystal structure is relatively high (1.56 of the final model). As follows from Tables 4.4 and 4.5, this is mainly caused by the high flexibility of the large loop regions. Indeed, the RMS deviation of the originally aligned regions from the corresponding C<sup>α</sup>-atoms in the crystal structure remains below 1.0. In spite of restraining the C<sup>α</sup>-atoms in these loop regions, the resulting movement is likely to have been caused by the absence of solvent. This especially accounts for the long loop connecting B3 and B4. This loop would have completely moved toward the bulk of the protein, if no restraints had been applied. As a consequence, B3 and B4 would have been distorted as well, which in turn would have affected the position of the structural water molecule. Since it was not found to be very useful to restrain this loop only, additional restraints were placed on the C<sup>α</sup>-carbon atoms of both β-sheets. Similarly, it was necessary to restrain the C<sup>α</sup> atoms in the large loop spanning residues T352 to F392, because otherwise movement of this region towards the bulk of the protein resulting from Van der Waals forces would have lowered the accessibility of the active site considerably. It has recently been confirmed that the side chains in this region are highly flexible and fold differently depending on the ligand bound [184].

## 5.4 Homology model of the transpeptidase domain of PBP2x\*

The Ramachandran plot presented in Figure 4.2 shows that 79.3 % of the 304 non-glycine and non-proline residues contained in the transpeptidase domain have most favoured geometry. Another 17.8 % occur in the additional allowed regions. Since only 3 of these residues - all located at the ends of long loops - have bad geometry, it can be concluded that the quality of the model is good. In fact, this is also reflected by the geometrical factors of the final model presented in Tables 4.6 and 4.7.

The good quality of the model was confirmed by Dr Dideberg [184] after comparison with the newly solved 2.4 Å resolution crystal structure [76]. Except for F359 and M386 all side chains are in the correct orientation. As expected, major differences occur in loop regions. The largest differences are represented by C<sup>α</sup><sub>480</sub> (3.4 Å) and C<sup>α</sup><sub>495</sub> (4.7 Å). The active site appears to be modelled correctly (except for the left side, but this is the region for which new coordinates had been supplied). The largest difference in this region occurs for C<sup>α</sup><sub>550</sub> (RMSD = 1.8 Å), but this is exactly where the loop connecting B3 and B4 starts.

Figure 4.3 shows the active site of the final model, including the structural water molecule forming hydrogen bonds with P335, T338, S371 and Y586. Even without knowing the location of additional structural water molecules in the active site, it can be concluded that the dense hydrogen bonding network present in the Class A β-lactamases is also present in PBP2x. The active site is nicely flanked by the indole ring system of W374, which does not seem to be stabilised by any hydrogen bond in particular, though. As confirmed by experimental data [184], a weak hydrogen bond is present between the side chains of H394 and D373, whereas R384 does not interact with any particular amino acid.

Summarising, it can be concluded that the procedure developed here performs very well for the construction of the active site region of a protein that has been solved to low resolution, and of which the C<sup>α</sup>-atom coordinates represent the only available structural data. Thanks to the application of restraining forces on some of the C<sup>α</sup>-carbon atoms and the heavy atoms of the amino acids for which new coordinates were supplied, the active site is still accessible. Major differences occur in the loops, but these could have been lowered if a higher restraining force had been applied.

---

This remains a matter of compromise, however, since this would also have affected the flexibility of these regions, which in turn would have decreased the conformational freedom of the aligned segments. Moreover, since it is well known that the quality of a model strongly depends on the quality of the alignment (which is also subject to the user's intuition [197]), it may be that a model can be improved by laboriously testing various alignments. Both this aspect and finding out which parts should be restrained represent the most laborious stages in the whole procedure.

## 5.5 Complexes of PBP2x\*<sub>R6</sub>

### 5.5.1 General remarks

After having developed a reliable model of the transpeptidase domain of PBP2x, the model was complexed with inhibitors. Michaelis complexes of four antibiotics (two penicillins and two cephalosporins) were studied with molecular dynamics simulations. In all simulations, the antibiotic was firmly flanked by W374, whereas S395 turned upward in all simulations, thus maintaining a strong hydrogen bond interaction with the carboxylate group of the antibiotic. Although the recently published crystal structures [191] show that the side chain is not turned upward, it might be hypothesised, that S395 changes its conformation upon formation of a Michaelis complex. This would imply that after the ring strain has been released upon ring opening, the carboxylate group would move away from S395, allowing for the original situation to be restored.

### 5.5.2 PBP2x\*<sub>R6</sub> complexed with benzylpenicillin and cefotaxime

First, benzylpenicillin and cefotaxime were docked into the active site with AutoDock. Both resulting complexes were subjected to a molecular dynamics simulation. The respective RMSD graphs presented in Figure 4.6 and Figure 4.6 clearly show that the starting configuration of the complex with cefotaxime was much better than that of benzylpenicillin. Although it should be kept in mind that the force field used in AutoDock differs from the CVFF. The fact that the active site is very narrow, however, might be another explanation for the observed displacements during the dynamics simulation: in spite of making use of the simulated annealing technique it appears to be very difficult to generate good Michaelis complexes with docking techniques. Nevertheless, both figures clearly show that a stable configuration is reached during the second half of the simulations. The position of cefotaxime remains even stable throughout the whole simulation.



The stability of the complex with benzylpenicillin after 100 ps simulation is also reflected in the course of the total interaction energy shown in Figure 4.1. As expected, Figure 4.2 shows that the decrease in total interaction energy after about 100 ps is mainly caused by electrostatic interaction forces. As shown in Figure 4.3 the resulting displacement of the antibiotic only leads to a small increase in Van der Waals energy. So, although the carboxylate group is highly attracted by the positive potential resulting from the charged nitrogen of K547, it becomes completely trapped. As a result, the dispersion energy only slightly increases after 100 ps, whereas the repulsion energy increases more significantly. Nevertheless, the RMSD graph shown in Figure 4.6 clearly shows that a stable situation has been reached after about 100 ps. As shown in Figure 4.8 the rather large RMSD value during the second half of the simulation is caused by the benzyl side chain adopting a different conformation.

The stability of the complex is further shown by the distance graphs presented in Figure 4.7. Indeed, Figures 4.19a and 4.19b clearly show that the lactam carbonyl oxygen atom of benzylpenicillin remains nicely fixed in the oxyanion hole, resulting in the electrophilic carbon atom to remain in the vicinity of the nucleophilic oxygen of the active serine (Figure 4.7c). The carboxylate group not only remains fixed by a strong salt bridge with K547 (Figure 4.7d), but also by a strong hydrogen bond with the side chain of T526 (Figure 4.7e). However, due to the  $sp^3$  hybridisation of C2 the carboxylate group is not able to form a hydrogen bond with the hydroxyl group of T550 (Figure 4.7f). Both because of the tilting caused by this hydrogen bonding interaction and because of the flexibility of the benzyl substituent, however, the amide group can not be fully stabilised by the side chain of N397 and the backbone carbonyl group of T550, however.

The situation is quite different in the case of cefotaxime. An energetically stable complex is immediately adopted after heating to physiological temperature (Figures 4.22 to 4.27). Moreover, the  $sp^2$  hybridisation of O2 allows the carboxylate group to strongly interact with the hydroxyl group of T550 via hydrogen bonding. On the other hand, T526 turns its polar side chain toward the structurally bound water molecule rather than trying to maintain hydrogen bond interaction with the carboxylate group of cefotaxime. Intuitively, this is the more stable conformation indeed, because trying to maintain this hydrogen bonding interaction would imply the methyl group to be turned toward the water molecule, which of course is energetically unfavourable. Instead, the methyl group is now turned toward the hydrophobic part of the dihydrothiazine

ring system, thereby allowing the hydroxyl group to stabilise the upwardly turned S395. Thus, while still allowing for a strong salt bridge formation with K547 (Figure 4.7d), the distance between the carboxylate group of cefotaxime and the T526 side chain gradually increases during the simulation, whereas the distance of the hydrogen bond with T550 is remains unaltered (Figure 4.7f), allowing for the lactam carbonyl oxygen atom to remain fixed in the oxyanion hole (Figures 4.28a and b). Compared to benzylpenicillin, however, anchoring of the oxime substituent allows for the amide group to be more stabilised by hydrogen bonding interactions with the side chain of N397 and the backbone carbonyl group of T550.

### 5.5.3 PBP2x\*<sub>R6</sub> complexed with oxacillin and cephalothin

Both the complexes with oxacillin and cephalothin were built out of the complexes of benzylpenicillin and cefotaxime, respectively. The course of the total interaction energy of oxacillin is similar to that of benzylpenicillin (Figure 4.1). Again, this seems to be caused by a significantly increased electrostatic interaction ( Figure 4.2), whereas the Van der Waals contribution to the total interaction energy remains more or less constant ( Figure 4.3). The stability of the complex after the first 40 ps is shown by the course of the RMSD (Figure 4.6), as well as the distance graphs presented in Figure 4.7. Like benzylpenicillin, the carboxylate group preferably interacts with T526 and S548 via hydrogen bond formation due to  $sp^3$  hybridisation of C2 (Figures 4.38 and 4.39) due to  $sp^3$  hybridisation of C2. Like benzylpenicillin, the amide group of oxacillin is not fully stabilised by N397 and T550.

Like cefotaxime, however, thanks to  $sp^2$  hybridisation of C2 is the carboxylate group of cephalothin still able to interact with the side chain hydroxyl group of T550 (Figures 4.47 and 4.48). As indicated on Figure 4.8 the increased RMSD value with respect to cefotaxime is caused by a much larger flexibility of the thiophene substituent in cephalothin. Indeed, due to the hydrophobic character of this substituent, this part of the antibiotic can not be stabilised by hydrogen bonding interactions. Although the dihydrothiazine ring system is still highly stabilised by hydrogen bonding interactions (Figure 4.7) with the protein, the larger flexibility of the thiophene substituent compared to benzylpenicillin apparently causes the lactam carbonyl oxygen to be

less regularly stabilised by the backbone NH of T550. This finding is in agreement with IR measurements [79] of *covalently bound* complexes.

#### 5.5.4 Comparison of the complexes with PBP2x\*R6

Since bond breaking and bond formation processes can not be described with force field methods (see section 3.1.3), these methods are not able to give a complete description of the acylation process, not in the least because force field calculations neglect several energy terms (including desolvation energy of both the ligand and the active site, and entropy changes). Since cephalosporins are known [198] to undergo elimination of the 3'-side chain after acyl-enzyme formation, it is even more clear that a complete description is impossible. Although the applicability of these terms is highly limited in the calculation of binding data [199-202], a few attempts trying to calculate the influence of water are known [203, 204]. Nevertheless, by studying the Michaelis complexes of on PBP2x complexed with various antibiotics it is still possible to draw some conclusions as to their affinities. For example, although the average distances of the intermolecular hydrogen bonding interactions with T550 in the oxyanion hole are shorter for both of the penicillins compared to the two cephalosporins (Table 4.10), the hydrogen bonds with the amidic nitrogen of S\*337 are shorter for the cephalosporins. As reflected in the average distances of the electrophilic carbons to the nucleophilic serine oxygen, this enables a shorter distance between these atoms. Thus, although only two members of both classes of antibiotics have been studied here, based on these distances it can be concluded that the acylation step proceeds faster for cephalosporins than for penicillins, as is indeed confirmed by experimental data (see Table 1.1 on page 29). Also when taking the dispersion energies (Table 4.9) into account, it is clear that both cephalosporins are the better inhibitors. Moreover, the average structure of PBP2x\* complexed with cefotaxime shows that a relatively large, hydrophilic R1 substituent containing an oxime group enables the amide group to be fully stabilised by hydrogen bonding interactions with N397 and T550.

As shown by the average structures of the complexes of PBP2x\* with benzylpenicillin and oxacillin, the carboxylate group of these penicillins forms a strong hydrogen bond

with the hydroxyl group of T526. Due to  $sp^3$  hybridisation of C2 this carboxylate group appears to be turned away from the hydroxyl group of T550 for strong hydrogen bonding interaction with this amino acid. The C2 atom in both of the cephalosporins is  $sp^2$  hybridised, however. As a consequence, the carboxylate group of these cephalosporins is still able to interact with T550 through strong hydrogen bond formation. This observation might explain binding data [61,78] showing that the affinity for cefotaxime strongly decreases upon mutation of T550 into alanine, whereas the T550→A mutant becomes hypersensitive for benzylpenicillin and oxacillin. The same reasoning may account for the natural T301→A mutation in *Streptomyces* sp. R61 DD-peptidase [205].

In order to maintain the hydrogen bond with the carboxylate group as well, T526 would be forced to turn its methyl group toward the structurally bound water molecule. Instead of trying to maintain a hydrogen bond with the carboxylate group, this amino acid turns its hydroxyl group in the direction of this water molecule, thereby stabilising S395 and enabling the methyl group hydrophobically interact with the dihydrothiazine ring system.

## 5.6 Complexes with PBP2x\* T526→S

### 5.6.1 Introduction

According to Hakenbeck et al. [82] the first amino acid that was mutated in PBP2x of the naturally occurring mutant strain C505 (T526→S, L403→F and Q458→K) was the T526→S mutation, which is also located in the active site. Since both L403 and Q458 occur in helices, it is likely that these mutations have a destabilising effect on the protein secondary structure. Moreover, since both of these mutations occur at a distance of approximately 14 Å from the active serine, these mutations do not interact with the antibiotic directly. However, T526 is still in the active site, making direct interaction with the inhibitor possible. In order to study the influences of the T526→S mutation on the formation of a Michaelis complex, MD simulations of the PBP2x\*<sub>T526→S</sub> mutant protein complexed with benzylpenicillin and cefotaxime have been performed here.

### 5.6.2 Comparison of the complexes with PBP2x\*<sub>T526→S</sub>

In contrast to the simulation with the R6 protein, the total interaction energy of the complex with benzylpenicillin immediately decreases after the annealing phase (Figure 4.1). Indeed, both the Van der Waals and electrostatic contributions to the total interaction energy remain constant hereafter (Figures 4.50 and 4.51), indicating that a stable complex has soon been reached after the annealing phase.

This is also reflected by the fact that contrasting to the simulation of this antibiotic with the R6 enzyme, both the dispersion and repulsion energies (Figures 4.52 and 4.53) remain more or less constant throughout the simulation. The stability of the complex is also shown by the course of the RMSD ( Figure 4.6), as well as the distance graphs (Figure 4.7).

Interestingly, S526 now bridges with the carboxylate group from the other side. As a result, the inhibitor is slightly tilted, thereby enabling hydrogen bonding interaction

with T550, which is not possible in the R6 enzyme. Still, the dispersion energies are the same (Table 4.10). Nevertheless, it could be hypothesised that this tilting of the ligand results in a loss of activity due to destabilisation of the Michaelis complex. T526 would not be able to do so, because this would imply the methyl group being turned toward the structural water molecule. In order to verify this hypothesis it would be highly desirable to study the binding behaviour of the single T526→S mutation with genetic engineering techniques.

The effect of the T526→S mutation is even more evident in the case of cefotaxime. Like in the complex of the mutant enzyme with benzylpenicillin, S526 now bridges from the other side with the carboxylate group. However, since the R1 substituent is not only much larger, but is also stabilised by hydrophilic interactions with other amino acids, the dihydrothiazine ring system is now forced away from the active serine. Although the distance between the reactive atom pairs is increased by only 0.04 Å (Table 4.10) and the complex seems to be as stable as in the case of the R6 enzyme (compare Figures 4.27 and 4.63), it is clear from Figure 4.7a that the lactam carbonyl group is much less stabilised in the oxyanion hole. Large fluctuations of the distance between the lactam carbonyl oxygen and the amidic hydrogen of the active serine is caused by the fact that the backbone in this region is not stabilised anymore. Therefore, although reliable binding data are not available, it can be predicted that the affinity of the T526→S mutant protein for cefotaxime will be strongly decreased. Based on steric considerations this hypothesis may be even generalised for lactam antibiotics containing a large hydrophilic R1 substituent.

## **6 Summary / Zusammenfassung**

---





The aim of this thesis was to first generate a full 3D model of the transpeptidase domain of *Streptococcus pneumoniae* penicillin-binding protein 2x (R6 strain) from its  $\alpha$ -carbon coordinates, to study the behaviour of penicillins and cephalosporins in the active site with molecular dynamics simulations of Michaelis complexes, and to study the influence of several natural mutations in the active site on the formation of these complexes.

A new method was developed for generating full protein coordinates from  $\alpha$ -carbon coordinates of a crystal structure solved to 3.5 Å resolution. Since after generation of full coordinates of the transpeptidase domain the structure had been solved to 2.4 Å resolution, new X-ray coordinates for the worst modeled loop (residues T370 to M386; 17 out of a total number of 351 residues constituting the transpeptidase domain) were incorporated, as kindly provided by Dr Dideberg.

The final model was complexed with benzylpenicillin, oxacillin, cefotaxime and cephalothin. By running molecular dynamics simulations of the complexes it could be predicted that lactam antibiotics with a large hydrophilic R1 substituent are more potent to covalently inhibit PBP2x than those with small hydrophobic substituents.

In addition, the molecular dynamics simulations of the complexes nicely explained experimental results, showing that the naturally occurring T550→A mutant enzyme is highly resistant to cefotaxime, but hypersensitive for benzylpenicillin and oxacillin.

Likewise, due to larger flexibility of the thiophene substituent compared to the benzyl substituent, simulations of the complexes could explain IR measurements, showing that the lactam carbonyl carbon of *covalently bound* cephalothin was less stabilised in the oxyanion hole than in covalently bound benzylpenicillin.

In order to study the effects of the T526→S mutation present in the naturally occurring mutant C505 strain (T526→S, L403→F and Q458→K), additional molecular dynamics simulations were run of this mutant enzyme complexed with benzylpenicillin and cefotaxime. Contrasting to the R6 enzyme, these simulations predicted that the larger the hydrophilic R1 substituent, the higher the resistance profile to be observed.

Im Rahmen der vorliegenden Arbeit wurde mit Hilfe von Molecular Modelling Methoden ein 3D Modell der Transpeptidase Domäne von *Streptococcus pneumoniae* Penicillin-bindendem Protein 2x (PBP2x) erstellt. Um mögliche Unterschiede im Verhalten der Penicilline und Cephalosporine im aktiven Zentrum aufzudecken, wurden Moleküldynamiken der Michaelis Komplexe durchgeführt.

Da zu Beginn nur die  $\alpha$ -Kohlenstoff-Koordinaten von einer bis zu 3,5 Å aufgelösten Kristallstruktur bekannt waren, wurde zuerst eine neue Methode zur Generierung von 3D Kompletstrukturen aus  $\alpha$ -Kohlenstoff Koordinaten entwickelt. Nachdem mit Hilfe dieser Methode ein komplettes 3D Modell der Transpeptidase-Domäne erstellt worden war, wurde die Struktur bis zu einer Auflösung von 2.4 Å gelöst. In das erstellte Modell brauchten nur die neuen Kristall-Koordinaten von 17 der 351 Aminosäuren eingefügt zu werden.

Das Endmodell wurde mit Benzylpenicillin, Oxacillin, Cefotaxim und Cefalothin komplexiert. Aus den Moleküldynamiksimulationen der Komplexe konnte vorhergesagt werden, dass Laktamantibiotika mit einem grossen hydrophilen R1 Substituenten PBP2x besser kovalent inhibieren können als Laktamantibiotika mit einem kleinen Substituenten. Ausserdem konnten die Simulationen experimentelle Befunde, laut denen die natürliche T550→A Mutante hoch resistent gegen Cefotaxim, aber hypersensitiv für Benzylpenicillin und Oxacillin ist, gut erläutern.

Da die Dynamiken zeigten, dass der Thiophensubstituent von Cefalothin beweglicher ist als der Benzylsubstituent von Benzylpenicillin, konnten IR Messungen, welche zeigen, dass der Lactam-Carbonylsauerstoff des *kovalent gebundenen* Cefalothin in der Oxyanion-Höhle weniger stabilisiert ist als im kovalenten Komplex mit Benzylpenicillin, ebenfalls erläutert werden. Zur Untersuchung der Effekte der in dem natürlichen mutanten Strang C505 vorkommenden T526→S Mutation, wurden zusätzliche Dynamiksimulationen der Komplexe dieses mutierten Enzyms, komplexiert mit Benzylpenicillin und Cefotaxim, durchgeführt. Im Gegensatz zu dem R6 Enzym, konnte anhand dieser Simulationen vorhergesagt werden, dass grosse hydrophile R1 Substituenten in Cephalosporinen zu einer stärkeren Resistenz führen.

## References

---



- 
- [1] Fleming, A.: *Br. J. Exp. Pathol.* 10 (1929) 226-236
- [2] Chain, E.B.; Florey, H.W.; Gardner, A.D.; Heatley, N.G.; Jennings, M.A.; Orr-Ewing, J. et al., Penicillin as a chemotherapeutic agent, *Lancet* 1 (1940) 226-228
- [3] Murray, R.G.E.; Steed, P.; Elson, H.E., *Can. J. Microbiol.*, 11 (1965) 547-560
- [4] Höltje, J.-V., Das unverstandene Penicillin, *Deutsche Apotheker Zeitung* 133 (1993) 505-513
- [5] Dijkstra, A.J.; Keck, W., Peptidoglycan as a barrier to transenvelope transport, *J. Bacteriol.* 178 (1996) 5555-5562
- [6] Weidel, W.; Pelzer, H., Bagshaped macromolecules: A new outlook on bacterial cell walls, *Adv. Enzymol.* 26 (1964) 193-232
- [7] Höltje, J.-V.; Schwarz, U., Biosynthesis and growth of the murein sacculus, In: *Molecular Cytology of Escherichia coli*, N. Nanninga, Ed., Academic Press, London 1985, pp. 77-119
- [8] Auterhoff, H.; Knabe, J.; Höltje, H.-D., *Lehrbuch der Pharmazeutischen Chemie*, 13. Aufl., Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart, 1994, p. 742
- [9] Labischinski, H.; Barnickel, G.; Bradaczek, H.; Giesbrecht, P., On the secondary and tertiary structure of murein. Low and medium-angle X-ray evidence against chitin-based conformations of bacterial peptidoglycan, *Eur. J. Biochem.* 95 (1979) 147-155
- [10] Thwaites, J.J.; Mendelson, N.H., Mechanical behaviour of bacterial cell walls, *Adv. Microb. Physiol.* 32 (1991) 173-222
- [11] Labischinski, H.; Barnickel, G.; Naumann, D., The state of order of bacterial peptidoglycan. In: *The target of penicillin*, R. Hakenbeck, J.-V. Höltje and H. Labischinski (Eds.), Walter de Gruyter, Berlin, 1983, pp. 49-54
- [12] van Heijenoort, J.: Biosynthesis of peptidoglycan. In: *New Comprehensive Biochemistry. The Bacterial Cell Wall (Vol. 27)*, J.-M. Ghuyesen and R. Hakenbeck (Eds.), Elsevier, Amsterdam, 1995, pp. 39-54
- [13] van Heijenoort, J., Assembly of the monomer unit of bacterial peptidoglycan, *Cell. Mol. Life Sci.* 54 (1998) 300-304
- [14] Matsushashi, M., Utilization of lipid-linked precursors and the formation of peptidoglycan in the process of cell growth and division: membrane enzymes involved in the final steps of peptidoglycan synthesis and the mechanism of their regulation, In: *New Comprehensive Biochemistry. Bacterial Cell Wall (Vol. 27)*, J.-M. Ghuyesen and R. Hakenbeck (Eds.), Elsevier, Amsterdam, 1994, pp. 55-71
- [15] Ghuyesen, J.-M., Serine beta-lactamases and penicillin-binding proteins, *Annu. Rev. Microbiol.* 45 (1991) 37-67
- [16] Ghuyesen, J.-M.; Frère, J.-M.; Leyh-Bouille, M.; Nguyen-Distèche, M.; Coyette, J., Active-site-serine D-alanyl-D-alanine-cleaving-peptidase-catalysed acyl-transfer reactions, *Biochem. J.* 235 (1986) 159-165
- [17] Frère, J.-M.; Joris, B., Penicillin-sensitive enzymes in peptidoglycan biosynthesis, *Crit. Rev. Microbiol.* 11 (1985) 299-396

- [18] Frère, J.-M.; Joris, B., Penicillin-sensitive enzymes in peptidoglycan synthesis, *CRC Crit. Rev. Microbiol.* 11 (1985) 299-396
- [19] Suginaka, H.; Blumberg, P.M.; Strominger, J.L., Multiple penicillin-binding proteins in *Bacillus subtilis*, *Bacillus cereus*, *Staphylococcus aureus* and *Escherichia coli*, *J. Biol. Chem.* 247 (1972) 5279-5288
- [20] Spratt, B.G.; Pardee, A.B., Penicillin-binding proteins and cell-shape in *Escherichia coli*, *Nature* 254 (1975) 516-517
- [21] Spratt, B.G.; Jobanputra, V.; Schwarz, U., Mutants of *Escherichia coli* which lack a component of penicillin-binding protein 1 are viable, *FEBS Lett.* 79 (1977) 374-378
- [22] Tamaki, S.; Nakajima, S.; Matsushashi, M., Thermosensitive mutation in *Escherichia coli* simultaneously causing defects in penicillin-binding proteins-1Bs and in enzyme activity for peptidoglycan synthesis in vitro, *Proc. Natl. Acad. Sci. USA* 74 (1977) 5472-5476
- [23] Suzuki, H.; Nishimura, Y.; Hirota, Y., On the process of cellular division in *Escherichia coli*: a series of mutants of *E. coli* altered in the penicillin-binding proteins, *Proc. Natl. Acad. Sci. USA* 75 (1978) 664-668
- [24] Spratt, B.G., Distinct penicillin-binding proteins involved in the division, elongation and shape of *Escherichia coli* K12, *Proc. Natl. Acad. Sci. USA* 72 (1975) 2999-3003
- [25] Hakenbeck, R.; Coyette, J., Resistant penicillin-binding proteins, *Cell. Mol. Life Sci.* 54 (1998) 332-340
- [26] Ghuysen, J.-M., Molecular structures of penicillin-binding proteins and  $\beta$ -lactamases, *Trends Microbiol.* 2 (1994) 372-380
- [27] Ghuysen, J.-M., Serine  $\beta$ -lactamases and penicillin-binding proteins, *Annu. Rev. Microbiol.* 45 (1991) 37-67
- [28] Ghuysen, J.-M.; Dive, G., Biochemistry of the penicilloyl transferases, In: *New Comprehensive Biochemistry. Bacterial Cell Wall* (Vol. 27), J.-M. Ghuysen and R. Hakenbeck (Eds.), Elsevier, Amsterdam, 1994, pp. 103-129
- [29] Englebert, S.; El Kharroubi, A.; Piras, G.; Joris, B.; Coyette, J.; Nguyen-Distèche, M.; Ghuysen, J.-M., Modular design of the bi(multi?)-functional penicillin-binding proteins, In: *Bacterial Growth and Lysis. Metabolism and structure of the bacterial sacculus* (FEMS Symposium No. 65), M.A. de Pedro, J.-V. Høltje and W. Löffelhardt (Eds.), Plenum Press, New York, 1993, pp. 319-333
- [30] Goffin, C; Ghuysen, J.-M., Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs, *Microbiol. Molec. Biol. Rev.* 62 (1998) 1079-1093
- [31] Zhu, Y.F.; Curran, I.J.; Joris, B; Ghuysen, J.-M.; Lampen, J.O., Identification of BlaR, signal transducer for  $\beta$ -lactamase production in *Bacillus licheniformis* as a penicillin-binding protein with strong homology to the Oxa-2  $\beta$ -lactamase (class D) of *Salmonella typhimurium*, *J. Bacteriol.* 172 (1990) 1137-1141

- [32] Joris, B.; Hardt, K.; Ghuysen, J.-M., Induction of  $\beta$ -lactamase and low-affinity penicillin-binding protein 2' synthesis in Gram-positive bacteria, In: *New Comprehensive Biochemistry. Bacterial Cell Wall* (Vol. 27), J.-M. Ghuysen and R. Hakenbeck (Eds.), Elsevier, Amsterdam, 1994, pp. 505-515
- [33] Tipper, D.J.; Strominger, J.L., Bacterial cell wall synthesis and structure in relation to the mechanism of action of penicillins and other antibacterial agents, *Proc. Natl. Acad. Sci. USA* 54 (1965) 1133-1141
- [34] Lee, B., Conformation of penicillin as a transition-state analog of the substrate of peptidoglycan transpeptidase, *J. Mol. Biol.* 61 (1971) 463-469
- [35] Frau, J.; Price, S.L., On the electrostatic and steric similarity of lactam compounds and the natural substrate for bacterial cell-wall biosynthesis, *J. Comp.-Aided Mol. Des.* 10 (1996) 107-122
- [36] <sup>a</sup>Frère, J.-M.; Ghuysen, J.-M.; Iwatsubo, M., Kinetics of the interaction between the exocellular DD-carboxypeptidase-transpeptidase from *Streptomyces* R39, *Eur. J. Biochem.* 57 (1975) 343-351 <sup>b</sup>Fuad, N.; Frère, J.-M.; Ghuysen, J.-M.; Duez, C.; Iwatsubo, M., Mode of interaction between beta-lactam antibiotics and the exocellular DD-carboxypeptidase-transpeptidase from *Streptomyces* R39 and beta-lactam antibiotics. A choice of models, *Biochem. J.* 155 (1976) 623-629
- [37] Marquet, A.; Frère, J.-M.; Ghuysen, J.-M.; Loffet, A., Effects of nucleophiles on the breakdown of benzylpenicilloyl-enzyme complex EI formed between benzylpenicillin and the exocellular DD-carboxypeptidase-transpeptidase of *Streptomyces* strain R61, *Biochem. J.* 177 (1979) 909-916
- [38] Adam, M.; Damblon, C.; Jamin, M.; Zorzi, W.; Dusart, V.; Galleni, M.; El Kharroubi, A.; Piras, G.; Spratt, B.G.; Keck, W.; Coyette, J.; Ghuysen, J.-M.; Frère, J.-M., Acyltransferase activities of the high-molecular-mass essential penicillin-binding proteins, *Biochem. J.* 279 (1991) 601-604
- [39] Adam, M.; Damblon, C.; Plaitin, B.; Christiaens, L.; Frère, J.-M., Chromogenic depsipeptide substrates for  $\beta$ -lactamases and penicillin-sensitive DD-peptidases, *Biochem. J.* 270 (1990) 525-529
- [40] Jamin, M.; Damblon, C.; Miller, S.; Hakenbeck, R.; Frère, J.-M., Penicillin-binding protein 2x of *Streptococcus pneumoniae*: enzymic activities and interactions with  $\beta$ -lactams, *Biochem. J.* 292 (1993) 735-741
- [41] Setti, E.L.; Micetrich, R.G., New trends in antimicrobial development, *Curr. Med. Chem.* 5 (1998) 101-113
- [42] Galleni, M.; Lamotte-Brasseur, J.; Raquet, X.; Dubus, A.; Monnaie, D.; Knox, J.R.; Frère, J.-M., The enigmatic catalytic mechanism of active-site serine  $\beta$ -lactamases, *Biochem. Pharmacol.* 49 (1995) 1171-1178

- [43] Bush, K.; Jacoby, G.A.; Medeiros, A.A., A functional classification scheme for  $\beta$ -lactamases and its correlation with molecular structure, *Antimicrob. Agents Chemother.* **39** (1995) 1211-1233
- [44] Ghuysen, J.-M.; Charlier, P.; Coyette, J.; Duez, C.; Fonzé, E.; Fraipont, C.; Goffin, C.; Joris, B.; Nguyen-Distèche, M., Penicillin and beyond: evolution, protein fold, multimodular polypeptides, and multiprotein complexes, *Microb. Drug Resist.* **2** (1996) 163-175
- [45] Ambler, R.P., The structure of  $\beta$ -lactamases, *Philos. Trans. R. Soc. London, B* **289** (1980) 321-331
- [46] Lamotte-Brasseur, J.; Dive, G.; Dideberg, O.; Charlier, P.; Frère, J.-M.; Ghuysen, J.-M., *Biochem J.* **279** (1991) 213-221
- [47] Bush, K., Characterization of beta-lactamases, *Antimicrob. Agents Chemother.* **33** (1989) 259-263
- [48] Massova, I; Mobashery, S., Structural and mechanistic properties of evolution of  $\beta$ -lactamases and penicillin-binding proteins, **5** (1999) 929-937
- [49] Kelly, J.A.; Dideberg, O.; Charlier, P.; Werz, J.P.; Libert, M.; Moews, P.C.; Knox, J.R.; Duez, C.; Fraipont, C.; Joris, B.; Dusart, J.; Frère, J.-M.; Ghuysen, J.-M., On the origin of bacterial resistance to penicillin: comparison of a beta-lactamase and a penicillin target, *Science* **231** (1986) 1429-1431
- [50] Samraoui, B.; Sutton, B.J.; Todd, R.J.; Artymiuk, P.J.; Waley, S.G.; Phillips, D.C., Tertiary structure similarity between a class A  $\beta$ -lactamase and a penicillin-sensitive D-alanyl carboxypeptidase-transpeptidase, *Nature* **320** (1986) 378-320
- [51] Varetto, L.; De Meester, F.; Monnaie, D.; Marchand-Brynaert, J.; Dive, G.; Jacob, F.; Frère, J.-M., The importance of the negative charge of  $\beta$ -lactam compounds in the interactions with active-site serine DD-peptidases and  $\beta$ -lactamases, *Biochem. J.* **278** (1991) 801-807
- [52] Bacquero, F., Pneumococcal resistance to  $\beta$ -lactam antibiotics, *Microb. Drug Resist.* **1** (1995) 115-120
- [53] Hakenbeck, R., Target mediated resistance to  $\beta$ -lactam antibiotics, *Biochem. Pharmacol.* **50** (1995) 1121-1127
- [54] Spratt, B.G., Resistance to  $\beta$ -lactam antibiotics, In: *New Comprehensive Biochemistry. Bacterial Cell Wall* (Vol. 27), J.-M. Ghuysen and R. Hakenbeck (Eds.), Elsevier, Amsterdam, 1994, pp. 517-534
- [55] Hakenbeck, R., Mosaic genes and their role in penicillin-resistant *Streptococcus pneumoniae*, *Electrophoresis* **19** (1998) 597-601
- [56] Tomasz, A., Benefit and risk in the  $\beta$ -lactam antibiotic-resistance strategies of *Streptococcus pneumoniae* and *Staphylococcus aureus*, *Trends Microbiol.* **2** (1994) 380-385
- [57] Hakenbeck, R.; Ellerbrok, H.; Briese, T., Antibodies against the benzylpenicillin moiety as a probe for penicillin-binding proteins, *Eur. J. Biochem.* **157** (1986) 101-106



- [58] Hakenbeck, R.; Kohiyama, M., Purification of penicillin-binding protein 3 from *Streptococcus pneumoniae*, *Eur. J. Biochem.* 127 (1982) 231-236
- [59] Severin, A.; Schuster, C.; Hakenbeck, R.; Tomasz, A., Altered murein composition in a DD-carboxypeptidase mutant of *Streptococcus pneumoniae*, *J. Bacteriol.* 174 (1992) 5152-5155
- [60] Dowson, C.G.; Johnson, A.P.; Cercenado, E.; George, R.C., Genetics of oxacillin-resistance in clinical isolates of *Streptococcus pneumoniae* that are oxacillin resistant and penicillin susceptible, *Antimicrob. Agents Chemother.* 38 (1994) 49-53
- [61] Grebe, T.; Hakenbeck, R., Penicillin-binding proteins 2b and 2x of *Streptococcus pneumoniae* are primary resistance determinants for different classes of  $\beta$ -lactam antibiotics, *Antimicrob. Agents Chemother.* 40 (1996) 829-834
- [62] Laible, G.; Hakenbeck, R.; Sicard, M.A.; Joris, B.; Ghuysen, J.-M., Nucleotide sequences of the *pbpX* genes encoding the penicillin-binding protein 2x from *Streptococcus pneumoniae* R6 and a cefotaxime-resistant mutant, C506, *Mol. Microbiol.* 3 (1989) 1337-1348
- [63] Kell, C.M.; Sharma, U.K.; Dowson, C.G.; Town, C.; Balganes, T.S.; Spratt, B.G., Deletion analysis of the essentiality of penicillin-binding proteins 1A, 2B and 2X of *Streptococcus pneumoniae*, *FEMS Microbiol. Lett.* 106 (1993) 171-176
- [64] Muñóz, R.; Dowson, C.G.; Daniels, M.; Coffey, T.J.; Martin, C.; Hakenbeck, R.; Spratt, B.G., Genetics of resistance to third-generation cephalosporins in clinical isolates of *Streptococcus pneumoniae*, *Mol. Microbiol.* 6 (1992) 2461-2465
- [65] Hakenbeck, R.; König, A.; Kern, I.; Van der Linden, M.; Keck, W.; Billot-Klein, D.; Legrand, R.; Schoot, B.; Gutmann, L., Acquisition of five high- $M_r$  penicillin-binding protein variants during transfer of high-level  $\beta$ -lactam resistance from *Streptococcus mitis* to *Streptococcus pneumoniae*, *J. Bacteriol.* 180 (1998) 1831-1840
- [66] Hakenbeck, R.; Grebe, T.; Zähler, D.; Stock, J.B.,  $\beta$ -Lactam resistance in *Streptococcus pneumoniae*: penicillin-binding proteins and non-penicillin-binding proteins, *Mol. Microbiol.* 33 (1999) 673-678
- [67] Hakenbeck, R.; Tornette, S.; Adkinson, N.F., Interaction of non-lytic  $\beta$ -lactams with penicillin-binding proteins in *Streptococcus pneumoniae*, *J. Gen. Microbiol.* 133 (1987) 755-760
- [68] Reichmann, P.; König, A.; Liñares, J.; Alcaide, F.; Tenover, F.C.; McDougal, L.; Swidsinski, S.; Hakenbeck, R., A global gene pool for high-level cephalosporin resistance in commensal *Streptococcus* spp. and *Streptococcus pneumoniae*, *J. Infect. Dis.* 176 (1997) 1001-1012
- [69] Laible, G.; Hakenbeck, R., Penicillin-binding proteins in  $\beta$ -Lactam resistant laboratory mutants of *Streptococcus pneumoniae*, *Mol. Microbiol.* 1 (1987) 355-363
- [70] Guenzi, E.; Gasc, A.M.; Sicard, M.A.; Hakenbeck, R., A two-component signal-transducing system is involved in competence and penicillin-susceptibility in laboratory mutants of *Streptococcus pneumoniae*, *Mol. Microbiol.* 12 (1994) 505-515

- [71] Grebe, T.; Paik, J.; Hakenbeck, R., A novel resistance mechanism for  $\beta$ -lactams in *Streptococcus pneumoniae* involves CpoA, a putative glycosyltransferase, *J. Bacteriol.* **179** (1997) 3342-3349
- [72] Laible, G.; Keck, W.; Lurz, R.; Mottl, H.; Frère, J.-M.; Jamin, M.; Hakenbeck, R., Penicillin-binding protein 2x of *Streptococcus pneumoniae*. Expression in *Escherichia coli* and purification of a soluble enzymatically active derivative, *Eur. J. Biochem.* **207** (1992) 943-949
- [73] Charlier, P.; Buisson, G.; Dideberg, O.; Wierenga, J.; Keck, W.; Laible, G.; Hakenbeck, R., Crystallization of a genetically engineered water-soluble primary penicillin target enzyme, *J. Mol. Biol.* **232** (1993) 1007-1009
- [74] Pares, S.; Mouz, N.; Pétilot, Y.; Hakenbeck, R.; Dideberg, O., X-ray structure of *Streptococcus pneumoniae* PBP2x, a primary penicillin target enzyme, *Nature Struct. Biol.* **3** (1996) 284-289
- [75] Protein Data Bank (URL: <http://www.rcsb.org/pdb/>)
- [76] Gordon, E.; Mouz, N.; Di Guilmi, A.-M.; Martin, L.; Duee, E.; Vernet, T.; Dideberg, O., Structural studies of penicillin binding protein PBP2x from *Streptococcus pneumoniae*, Sectoral Meeting: „Controlling the Proliferation of the Microbial Cell Factory“, Verona, Italy, April 19-21, 1999
- [77] Mouz, N.; Gordon, E.; Di Guilmi, A.-M.; Petit, I.; Petillot, Y.; Dupont, Y.; Hakenbeck, R.; Vernet, T.; Dideberg, O., Identification of a structural determinant for resistance to  $\beta$ -lactam antibiotics in Gram-positive bacteria, *Proc. Natl. Acad. Sci. USA* **95** (1998) 13403-13406
- [78] Mouz, N.; Di Guilmi, A.-M.; Gordon, E.; Hakenbeck, R.; Dideberg, O.; Vernet, T., Mutations in the active site of penicillin-binding protein PBP2x from *Streptococcus pneumoniae*, *J. Biol. Chem.* **274** (1999) 19175-19180
- [79] Chittock, R.S.; Ward, S.; Wilkinson, A.-S.; Caspers, P.; Mensch, B.; Page, M.G.P.; Wharton, C.W., Hydrogen bonding and protein perturbation in  $\beta$ -lactam acyl-enzymes of *Streptococcus pneumoniae* penicillin-binding protein PBP2x, *Biochem. J.* **338** (1999) 153-159
- [80] Laible, G.; Hakenbeck, R., Five independent combinations of mutations can result in low-affinity penicillin-binding protein 2x of *Streptococcus pneumoniae*, *J. Bacteriol.* **173** (1991) 6986-6990
- [81] Asahi, Y.; Takeuchi, Y.; Ubukata, K., Diversity of substitutions within or adjacent to conserved amino acid motifs of penicillin-binding protein 2X in cephalosporin-resistant *Streptococcus pneumoniae* isolates, *Antimicrob. Agents Chemother.* **43** (1999) 1252-1255
- [82] Krauß, J.; van der Linden, M.; Frère, J.-M.; Gordon, E.; Dideberg, O.; Hakenbeck, R., unpublished results
- [83] Garcia-Bustos, J.; Tomasz, A., A biological price of antibiotic resistance: major changes in the peptidoglycan structure of penicillin-resistant pneumococci, *Proc. Natl. Acad. Sci. USA* **87** (1990) 5414-5419

- [84] Severin, A.; Figueiredo, A.M.S.; Tomasz, A., Separation of abnormal cell wall composition from penicillin resistance through genetic transformation of *Streptococcus pneumoniae*, *J. Bacteriol.* 178 (1996) 1788-1792
- [85] Severin, A.; Tomasz, A., Naturally occurring peptidoglycan variants of *Streptococcus pneumoniae*, *J. Bacteriol.* 178 (1996) 168-174
- [86] Severin, A.; Vaz Pato, M.V.; Figueiredo, A.M.S.; Tomasz, A., Drastic changes in the peptidoglycan composition of penicillin resistant laboratory mutants of *Streptococcus pneumoniae*, *FEMS Microbiol. Lett.* 130 (1995) 31-35
- [87] Zhao, G.; Yeh, W.-K.; Carnahan, R.H.; Flokowitsch, J.; Meier, T.I.; Alborn, W.E.; Becker, G.W.; Jaskunas, S.R., Biochemical characterization of penicillin-resistant and -sensitive penicillin-binding protein 2x transpeptidase activities of *Streptococcus pneumoniae* and mechanistic implications in bacterial resistance to  $\beta$ -lactam antibiotics, *J. Bacteriol.* 179 (1997) 4901-4908
- [88] Boulnois, G.J., Pneumococcal proteins and the pathogenesis of disease caused by *Streptococcus pneumoniae*, *J. Gen. Microbiol.* 138 (1992) 249-259
- [89] Roberts, R.B., *Streptococcus pneumoniae*, In: Principles and practices of infectious diseases, G.L. Mandell, R.G. Douglas, J.E. Bennett (Eds.), Churchill Livingstone, New York, 1985, pp. 1142-1152
- [90] Austrian, R., Some aspects of the pneumococcal carrier state, *J. Antimicrob. Chemother.* 18 (1986) 35-45
- [91] Felmingham, D.; Grunberg, R.N., The Alexander Project 1996-1997: latest susceptibility data from this international study of bacterial pathogens from community-acquired lower respiratory tract infections, *J. Antimicrob. Chemother.* 45 (2000) 191-203
- [92] Hermans, P.W.; Sluiter, M.; Elzenaar, K.; van Veen, A.; Schonkeren, J.J.; Nooren, F.M.; van Leeuwen, W.J.; de Neeling, A.J.; van Klingeren, B.; Verbrugh, H.A.; de Groot, R., Penicillin-resistant *Streptococcus pneumoniae* in the Netherlands: results of a 1-year molecular epidemiologic survey, *J. Infect. Dis.* 175 (1997) 1413-1422
- [93] Zbinden, R., Die verminderte Penicillin-empfindlichkeit der Pneumokokken, *Ther. Umsch.* 55 (1998) 18-21
- [94] Levine, I.N., Quantum Chemistry, Fifth Edition, Prentice Hall, London 1991, pp. 366ff.
- [95] Roothaan, C.C.J., New developments in Molecular Orbital Theory, *Rev. Mod. Phys.* 23 (1951) 69
- [96] Atkins, P.W.; Friedman, R.S., Molecular quantum mechanics, Third Edition, Oxford University Press, Oxford, 1997, pp. 285ff
- [97] Hartree, D.R., The Calculation of Atomic Structures, Wiley, New York 1957
- [98] Møller, C.; Plesset, M.S., *Phys. Rev.* 46 (1934) 618
- [99] Hoffmann, R., An extended Hückel theory. I. Hydrocarbons, *J. Chem. Phys.* 39 (1963) 1397-1412

- [100] Pople, J.A.; Segal, G.A., Approximate self-consistent molecular orbital theory. II. Calculations with complete neglect of differential overlap, *J. Chem. Phys.* **43** (1965) S136-S149
- [101] Pople, J.A.; Santry, D.P.; Segal, G.A., Approximate self-consistent molecular orbital theory. I. Invariant procedures, *J. Chem. Phys.* **43** (1965) S129-S135
- [102] Schmidtke, H.H., Quantenchemie, 2. Auflage, VCH Verlagsgesellschaft mbH, Weinheim, 1994, p. 254
- [103] Baird, N.C.; Dewar, M.J.S., Ground states of  $\sigma$ -bonded molecules. IV. The MINDO method and its application to hydrocarbons, *J. Chem. Phys.* **50** (1969) 1262-1274
- [104] <sup>a</sup>Dewar, M.J.S.; Thiel, W., Ground states of molecules. 38. The MNDO method. Approximations and parameters, *J. Amer. Chem. Soc.* **99** (1977) 4899-4907 <sup>b</sup>Dewar, M.J.S.; Thiel, W., Ground states of molecules. 39. MNDO results for molecules containing hydrogen, carbon, nitrogen and oxygen, *J. Amer. Chem. Soc.* **99** (1977) 4907-4917 <sup>c</sup>Dewar, M.J.S.; Thiel, W., A semiempirical model for the two-center repulsion integrals in the NDDO approximation, *Theoret. Chim. Acta* **46** (1977) 89-104
- [105] Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.P., AM1: A new general purpose quantum mechanical molecular model, *J. Amer. Chem. Soc.* **107** (1985) 3902-3909
- [106] <sup>a</sup>Stewart, J.J.P., Optimization of parameters for semiempirical methods. I. Method, *J. Comput. Chem.* **10** (1989) 209-220, <sup>b</sup>Stewart, J.J.P., Optimization of parameters for semiempirical methods. II. Applications, *J. Comput. Chem.* **10** (1989) 221-264
- [107] Dinur, U.; Hagler, A.T., New approaches to empirical force fields, In: Reviews in Computational Chemistry, Vol. 2, K.B. Lipkowitz and D.B. Boyd (Eds.), VCH Publishers, Inc. New York, 1991, p. 106
- [108] <sup>a</sup>Hagler, A.T.; Huler, E.; Lifson, S., Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystal properties, *J. Am. Chem. Soc.* **96** (1974) 5319-5327 <sup>b</sup>Hagler, A.T.; Lifson, S., Energy functions for peptides and proteins. II. The amide hydrogen bond and calculation of amide crystal properties, *J. Am. Chem. Soc.* **96** (1974) 5327-5335
- [109] Morse, P.M., Diatomic molecules according to the wave mechanics. II. Vibrational levels, *Phys. Rev.* (1929) 57-64
- [110] Buckingham, A.D.; Utting, B.D., Intermolecular forces, *Annu. Rev. Phys. Chem.* **21** (1970) 287
- [111] Scheiner, S., Calculating the properties of hydrogen bonds by *ab initio* methods, In: Reviews in Computational Chemistry, Vol. 2, K.B. Lipkowitz and D.B. Boyd (Eds.), VCH Publishers, Inc. New York, 1991, pp.165-218
- [112] Hölting, H.-D.; Folkers, G., Molecular Modeling – Basic Principles and Applications. In: Methods and Principles in Medicinal Chemistry, Vol. 5, Mannhold, R.; Kubinyi, H.; Timmerman, H. (Eds.) VCH Weinheim, 1996, p. 38
- [113] Dinur, U.; Hagler, A.T., Geometry-dependent atomic charges: Methodology and application to alkanes, aldehydes, ketones and amides, *J. Comput. Chem.* **16** (1995) 154-170

- [114] Insight II, Molecular Simulations Incorporated, Version 97.0, Biosym/MSI, 9685 Scranton Road, San Diego, CA 92121-3752, USA
- [115] van Gunsteren, W.F.; Berendsen, J.C., Computer simulation of molecular dynamics: methodology, applications and perspectives in chemistry, *Angew. Chem. Int. Ed. Engl.* **29** (1990) 992-1023
- [116] Atkins, P.W., Physical Chemistry, Third Edition, Oxford University Press, Oxford 1987, Statistical Thermodynamics: the concepts, pp. 506-526
- [117] Bash, P.A.; Singh, U.C.; Kollman, P.A., Free energy calculations by computer simulation, *Science* **236** (1987) 564-568
- [118] Verlet, L., Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules, *Phys. Rev.* **159** (1967) 98-103
- [119] Moulton, J., Predicting protein three-dimensional structure, *Curr. Opin. Biotechnol.* **10** (1999) 583-588
- [120] Holm, L.; Sander, C., Mapping the protein universe, *Science* **273** (1996) 595-602
- [121] Clothia, C.; Lesk, A.M., Relationship between the divergence of sequence and structure in proteins, *EMBO J.* **5** (1986) 823-827
- [122] Sánchez, R.; Šali, A., Advances in comparative protein-structure modelling, *Curr. Opin. Struct. Biol.* **7** (1997) 206-214
- [123] Clothia, C., One thousand families for the molecular biologist, *Nature* **360** (1992) 543-544
- [124] Orengo, C.A.; Jones, D.T.; Thornton, J.M., Protein superfamilies and domain superfolds, *Nature* **372** (1994) 631-634
- [125] Šali, A., Modelling mutations and homologous proteins, *Curr. Opin. Biotechnol.* **6** (1995) 437-451
- [126] Šali, A., Comparative protein modelling by satisfaction of spatial restraints, *Mol. Med. Today* **1** (1995) 270-277
- [127] Doolittle, R.F., Convergent evolution: the need to be explicit, *Trends Biochem. Sci.* **19** (1994) 15-18
- [128] Sander, C.; Schneider, R., Database of homology-derived structures and the structural meaning of sequence alignment, *Proteins* **9** (1991) 56-68
- [129] Holm, L.; Sander, C., The FSSP database of structurally aligned protein fold families, *Nucleic Acids Res.* **22** (1994) 3600-3609
- [130] Fitch, W.M., An improved method of testing for evolutionary homology, *J. Mol. Biol.* **16** (1966) 9-16
- [131] Cantor, C.R.; Jukes, T.H., The repetition of homologous sequences in the polypeptide chains of certain cytochromes and globins, *Proc. Natl. Acad. Sci. USA* **56** (1966) 177-184
- [132] Needleman, S.B.; Wunsch, C.D., A general method applicable to the search for similarities in the amino acid sequence of two proteins, *J. Mol. Biol.* **48** (1970) 443-453

- [133] Murata, M, Richardson, J.S.; Sussman, J.L., Simultaneous comparison of three protein sequences, *Proc. Natl. Acad. Sci. USA* 82 (1985) 3073-3077
- [134] Smith, T.F.; Waterman, M.S., Identification of common molecular subsequences, *J. Mol. Biol.* 147 (1981) 195-197
- [135] Pearson, W.R.; Lipman, D.J., Improved tools for biological sequence comparison, *Proc. Natl. Acad. Sci. USA* 85 (1988) 2444-2448 (URL: <http://www.ebi.ac.uk/searches/fasta.html>)
- [136] Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J., Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403-410 (URL: <http://www.ncbi.nlm.nih.gov>)
- [137] Dayhoff, M.O.; Schwarz, R.M.; Orcutt, B.C., In: Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC, 1978, Vol. 5 Suppl. 3, pp. 345-352
- [138] Gonnet, G.H.; Cohen, M.A.; Benner, S.A., Exhaustive matching of the entire protein sequence database, *Science* 256 (1992) 1443-1445
- [139] Jones, D.T.; Taylor, W.R.; Thornton, T.M., The rapid generation of mutation data matrices from protein sequences, *Comput. Appl. Biosci.* 8 (1992) 275-282
- [140] Risler, J.L.; Delome, M.O.; Delacroix, H.; Henaut, A., Amino acid substitutions in structurally related proteins. A pattern recognition approach. Determination of a new and efficient scoring matrix, *J. Mol. Biol.* 204 (1988) 1019-1029
- [141] Johnson, M.S.; Overington, J.P., A structural basis for sequence comparisons. An evaluation of scoring methodologies, *J. Mol. Biol.* 233 (1993) 716-738
- [142] Lüthy, R.; McLachlan, A.D.; Eisenberg, D., Secondary structure-based profiles: use of structure-conserving scoring tables in searching protein sequence databases for structural similarities, *Proteins* 10 (1991) 229-239
- [143] Henikhoff, S.; Henikhoff, J.G., Amino acid substitution matrices from protein blocks, *Proc. Natl. Acad. Sci. USA* 89 (1992) 10915-10919
- [144] Tomii, K.; Kanehisa, M., Analysis of amino acid indices and mutation matrices for sequence comparison and structure prediction of proteins, *Protein Eng.* 9 (1996) 27-36
- [145] Browne, W.J.; North, A.C.T.; Phillips, D.C.; Brew, K.; Vanaman, T.C.; Hill, R.C., A possible three-dimensional structure of bovine  $\alpha$ -lactalbumin based on that of hen's egg-white lysozyme, *J. Mol. Biol.* 42 (1969) 65-86
- [146] Greer, J., Comparative model-building of the mammalian serine proteases, *J. Mol. Biol.* 153 (1981) 1027-1042
- [147] Blundell, T.L.; Sibanda, B.L.; Sternberg, M.J.E.; Thornton, J.M., Knowledge-based prediction of protein structures and the design of novel molecules, *Nature* 326 (1987) 347-352
- [148] Murzin, A.G.; Brenner, S.E.; Hubbard, T.; Clothia, C., SCOP: A Structural Classification of Proteins database for the investigation of sequences and structures, *J. Mol. Biol.* 247 (1995) 536-540

- [149] Jones, T.H.; Thirup, S., Using known substructures in protein model building and crystallography, *EMBO J.* 5 (1986) 819-822
- [150] Reid, L.S.; Thornton, J.M., Rebuilding flavodoxin from C alpha coordinates: a test study, *Proteins Struct.Funct. Genet.* 5 (1989) 170-182
- [151] Holm, L.; Sander, C., Database algorithm for generating protein backbone and side-chain coordinates from a C alpha trace application to model building and detection of co-ordinate errors, *J. Mol. Biol.* 218 (1991) 183-194
- [152] Unger, R.; Harel, D.; Wherland, W.; Sussman, J.L., A 3-D building blocks approach to analyzing and predicting structure of proteins, *Proteins* 5 (1989) 355-373
- [153] Claessens, M, van Cutsem, E.; Lasters, I.; Wodak, S., Modelling the polypeptide backbone with 'spare parts' from known protein structures, *Protein Eng.* 4 (1989) 335-345
- [154] Levitt, M., Accurate modeling of protein conformation by automatic segment matching, *J. Mol. Biol.* 226 (1992) 507-533
- [155] Šali, A.; Blundell, T., Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779-815
- [156] Havel, T.; Snow, M., A new method for building protein conformations from sequence alignments with homologues of known structure, *J. Mol. Biol.* 217 (1991) 1-7
- [157] Holm, L.; Sander, C., Protein structure comparison by alignment of distance matrices, *J. Mol. Biol.* 233 (1993) 123-138
- [158] Snow, M., A novel parametrization scheme for energy equations and its use to calculate the structure of protein molecules, *Proteins* 15 (1993) 183-190
- [159] <sup>a</sup> Schäfer, L.; Cao, M.; Meadows, M.J., Predictions of protein backbone bond distances and angles from first principles, *Biopolymers* 35 (1995) 603-606 <sup>b</sup> Schäfer, L.; Cao, M., Predictions of protein backbone bond distances and angles from first principles, *J. Mol. Struct.* 333 (1995) 201-208 <sup>c</sup> Jiang, X.; Cao, M.; Teppen, B.; Newton, S.Q.; Schäfer, L., Predictions of protein backbone structural parameters from first principles: systematic comparisons of calculated N-C( $\alpha$ )-C' angles with high resolution protein crystallographic results, *J. Phys. Chem.* 99 (1995) 10521-10525 <sup>d</sup> Jiang, X.; Cao, M.; Newton, S.Q.; Schäfer, L.; Paulus, E.F., Predictions of peptide and protein backbone structural parameters from first principles. IV: Systematic comparisons of calculated N-C( $\alpha$ )-C' angles with peptide crystal structures, *Electronic J. Theor. Chem.* 1 (1995) 11-17
- [160] <sup>a</sup> Gan, K.; Alexander, P.; Coxon, J.M.; McKinnon, A.J.; Worth, G.H., The reconstruction of a protein backbone from C $\alpha$  coordinates, *Biopolymers* 41 (1997) 381 <sup>b</sup> Gan, K.; Alexander, P.; Coxon, J.M.; McKinnon, A.J.; Worth, G.H., Assignment of secondary structure from C $\alpha$  coordinates, *Biopolymers* 41 (1997) 367
- [161] Milik, M.; Kolinski, A.; Skolnick, J., Algorithm for rapid reconstruction of protein backbone from alpha carbon coordinates, *J. Comp. Chem.* 18 (1997) 80-85
- [162] Purisima, E.O.; Scheraga, H.A., Conversion from a virtual-bond chain to a complete polypeptide backbone chain, *Biopolymers* 23 (1984) 1207-1224

- [163] <sup>a</sup> Rackovsky, S.; Scheraga, H.A., Influence of ordered backbone structure on protein folding. A study of some simple models, *Macromolecules* 11 (1978) 1168-1174 <sup>b</sup> Rackovsky, S.; Scheraga, H.A., Differential geometry and polymer conformation. 2. Development of a conformational distance function, *Macromolecules* 13 (1980) 1440-1453
- [164] <sup>a</sup> Wako, H.; Scheraga, H.A., Distance constraint approach to protein folding. I. Statistical analysis of protein conformations in terms of distances between residues, *J. Protein Chem.* 1 (1982) 5-45, <sup>b</sup> Wako, H.; Scheraga, H.A., Distance constraint approach to protein folding. II. Prediction of three-dimensional structure of bovine pancreatic trypsin inhibitor, *J. Protein Chem.* 1 (1982) 85-117
- [165] Liwo, A.; Pincus, M.R.; Wawak, R.J.; Rackovsky, S.; Scheraga, H.A., Calculation of backbone geometry from  $\alpha$ -carbon coordinates based on peptide group dipole alignment, *Protein Sci.* 2 (1993) 1697-1714
- [166] Nishikawa, K.; Momany, F.A.; Scheraga, H.A., Low-energy structures of two dipeptides and their relationship to bend conformations, *Macromolecules* 7 (1974) 797-806
- [167] Dill, K.A.; Dominant forces in protein folding, *Biochemistry* 29 (1990) 7133-7155
- [168] Rosenfeld, R.; Vajda, S.; DeLisi, C., Flexible docking and design, *Annu. Rev. Biophys. Biomol. Struct.* 24 (1995) 677-700
- [169] Kuntz, I.D., Structure-based strategies for drug design and discovery, *Science* 257 (1992) 1078-1082
- [170] Najmanovich, R.; Kuttner, J.; Sobolev, V.; Edelman, M., Side-chain flexibility in proteins upon ligand binding, *Proteins Struct. Funct. Genet.* 39 (2000) 261-268
- [171] DesJarlais, R.L.; Sheridan, R.P.; Seibel, G.L.; Dixon, J.S.; Kuntz, I.D.; Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure, *J. Med. Chem.* 31 (1988) 722-729
- [172] Shoichet, B.K.; Kuntz, I.D., Protein docking and complexity, *J. Mol. Biol.* 221 (1991) 327-346
- [173] Morris, G.M.; Goodsell, D.S.; Huey, R.; Olson, A.J., Distributed automated docking of flexible ligands to proteins: parallel applications of Autodock 2.4, *J. Comp.-Aided Mol. Design* 10 (1996) 293-304
- [174] Weiner, S.J.; Kollman, P.A.; Case, D.A.; Singh, U.C.; Ghio, C.; Alagona, G.; Profeta Jr., S.; Weiner, P., A new force field for molecular mechanical simulation of nucleic acids and proteins, *J. Am. Chem. Soc.* 106 (1984) 765-784
- [175] AutoDock manual (URL: <http://www.scripps.edu/pub/olson-web/doc/autodock/>)
- [176] Goodford, P.J., A computational procedure for determining energetically favorable binding sites on biologically important macromolecules, *J. Med. Chem.* 28 (1985) 849-857
- [177] Mehler, E.L.; Solmajer, T., Electrostatic effects in proteins: comparison of dielectric and charge models, *Protein Eng.* 4 (1991) 903-910



- [178] Hooft, P.A.M. van; Höltje, H.-D., Construction of a full three-dimensional model of the transpeptidase domain of *Streptococcus pneumoniae* PBP2x starting from its C<sup>α</sup>-atom coordinates, *J. Comp.-Aided Mol. Design* **14** (2000) 719-730
- [179] Payne, P.W., Reconstruction of protein conformations from estimated positions of the C alpha coordinates, *Protein Sci.* **2** (1993) 315-324
- [180] Dayhoff, M.O., Barker, W.C. and Hunt, L.T., Establishing homologies in protein sequences, *Methods Enzymol.* **91** (1983) 524-545
- [181] Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M., *J. Appl. Cryst.* **26** (1993) 283-291
- [182] Mouz, N.; Gordon, E.; DiGuilmi, A.-M.; Petit, I.; Petillot, Y.; Dupont, Y.; Hakenbeck, R.; Vernet, T.; Dideberg, O., Identification of a structural determinant for resistance to beta-lactam antibiotics in Gram-positive bacteria, *Proc. Natl. Acad. Sci. USA* **95** (1998) 13403-13406
- [183] Gordon, E.; Mouz, N.; Di Guilmi, A.M.; Martin, L.; Duee, E.; Vernet, T.; Dideberg, O., Sectoral Meeting: „Controlling the Proliferation of Microbial Cell Factory“, Verona, Italy, April 19-21,1999
- [184] Dideberg, O., personal communication
- [185] Engh, R.A.; Huber, R., *Acta Cryst.* **47** (1991) 392-400
- [186] Mosimann, S.; Meleshko, R.; James, M.N., A critical assessment of comparative molecular modeling of tertiary structures of proteins, *Proteins Struct. Funct. Genet.* **23** (1995) 301-317
- [187] SURFNET manual (URL: <http://www.pharma.ethz.ch/~didier/manuals/surfnet/man1.html>)
- [188] Gaedt, K., Theoretische Untersuchungen der Bindungsstelle der Dopamin D3 Rezeptor Agonisten, Düsseldorf, Heinrich-Heine-Universität Düsseldorf, Fachbereich Pharmazeutische Chemie, dissertation, 1997
- [189] Paulin, G., Molecular Modelling Untersuchungen zur Substratspezifität der Breitspektrum-β-Lactamase MEN1 aus *Escherichia coli*, Düsseldorf, Heinrich-Heine-Universität Düsseldorf, Fachbereich Pharmazeutische Chemie, dissertation, 1997
- [190] SYBYL 6.4, Tripos Associates, St. Louis, Missouri, USA
- [191] Gordon, E.; Mouz, N.; Duee, E.; Dideberg, O., The crystal structure of the penicillin-binding protein 2x from *Streptococcus pneumoniae* and its acyl-enzyme form: implication in drug resistance, *J. Mol Biol.* **299** (2000) 477-485
- [192] Correa, P.E., The building of protein structures from alpha-carbon coordinates, *Proteins* **7** (1990) 366-377
- [193] Rey, A.; Skolnick, J., Efficient algorithm for the construction of a protein backbone from the α-carbon coordinates, *J. Comput. Chem.* **13** (1992) 443-456
- [194] Moulton, J.; Hubbard, T., Fidelis, K.; Pedersen, J.T., Critical assessment of methods of protein structure prediction (CASP): round III, *Proteins Struct. Funct. Genet. Suppl.* **3** (1999) 2-6

- [195] Levy, R.M.; Srinivasan, A.R.; Olson, W.K.; McCammon, J.A., Quasi-harmonic method for studying very low frequency modes in proteins, *Biopolymers* 23 (1984) 1099-1112
- [196] Higo, J.; Umeyama, H., Protein dynamics determined by backbone conformation and atom packing, *Protein Eng.* 10 (1997) 373-380
- [197] Bates, P.A.; Sternberg, M.J.E., Model building by comparison at CASP3: Using expert knowledge and computer automation, *Proteins Struct.Funct. Genet.*, 37 (Suppl. 3) (1999) 47-54
- [198] Faraci, W.S.; Pratt, R.F., Mechanism of inhibition of the PC1 beta-lactamase of *Staphylococcus aureus* by cephalosporins: importance of the 3'-leaving group, *Biochemistry* 24 (1985) 903-910
- [199] Dunitz, J.D., The entropic cost of bound water in crystals and biomolecules, *Science* 264 (1994) 670
- [200] Bhat, T.N.; Bentley, G.A.; Boulot, G.; Greene, M.I.; Tello, D.; Dell' Aqua, W.; Souchon, H.; Schwarz, F.P.; Mariuzza, R.A.; Poljak, R.J., Bound water molecules and conformational stabilization help mediate an antigen-antibody association, *Proc. Natl. Acad. Sci. USA* 91 (1994) 1089-1093
- [201] Tame, J.R.H.; Murshudov, G.N.; Dodson, E.J.; Neil, T.K.; Dodson, G.G.; Higgins, C.F.; Wilkinson, A.J., The structural basis of sequence-dependent peptide binding by OppA protein, *Science* 264 (1994) 1578-1581
- [202] Pearlman, C.S.; Connelly, P.R., Determination of the differential effects of hydrogen bonding and water release on the binding of FK506 to native and Tyr82→Phe82 FKBP-12 proteins using free energy simulations, *J. Mol. Biol.* 248 (1995) 696-717
- [203] Tame, J.R.H., Scoring functions: a view from the bench, *J. Comput.-Aided Mol. Design* 13 (1999) 99-108
- [204] Jain, A.N., Scoring non-covalent protein-ligand interactions: A continuous differentiable function tuned to compute binding affinities, *J. Comput.-Aided Mol. Design* 10 (1996) 427-440
- [205] Kuzin, A.P.; Liu, H.; Kelly, J.A.; Knox, J.R., Binding of cephalothin and cefotaxime to D-Ala-D-Ala-peptidase reveals a functional base of a natural mutation in a low-affinity penicillin-binding protein and in extended-spectrum  $\beta$ -lactamases, *Biochemistry* 34 (1995) 9532-9540

## **Appendices**

---

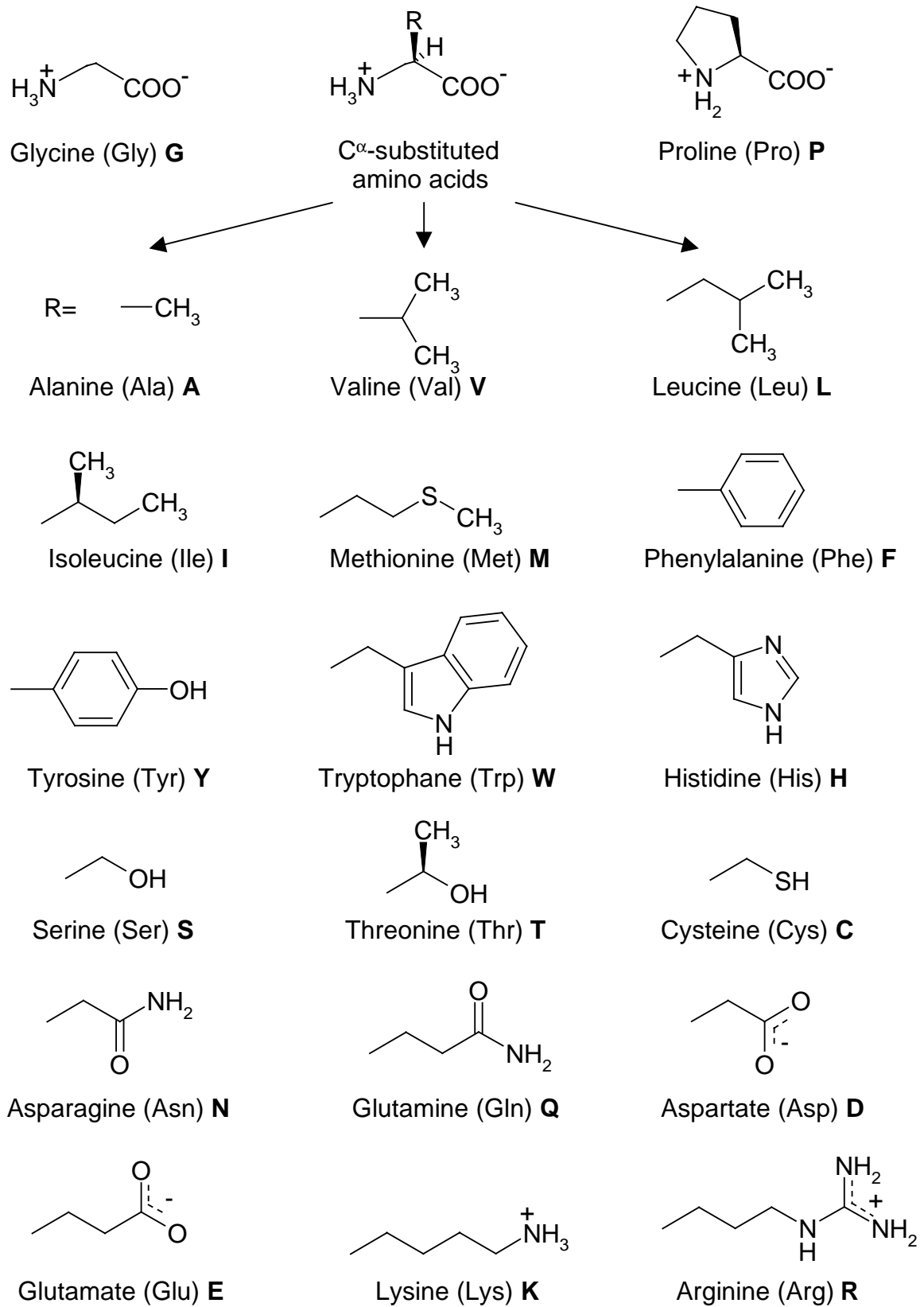


**Appendix A      List of abbreviations**

3D	Three-dimensional
A <sub>2</sub> pm	Diaminopimelic acid
AM1	Austin Model 1
AO	Atomic Orbital
BLAST	Basic Local Alignment Search Tool
CI	Configuration Interaction
CNDO	Complete Neglect of Differential Overlap
CSD	Cambridge Structural Database
CVFF	Consistent Valence Force Field
GlcNAc	<i>N</i> -acetylglucosamine
GTO	Gaussian Type Orbital
HF	Hartree-Fock
INDO	Intermediate Neglect of Differential Overlap
LCAO	Linear Combination of Atomic Orbitals
MD	Molecular Dynamics
MIC	Minimal Inhibition Concentration
MINDO	Modified Intermediate Neglect of Differential Overlap
MM	Molecular Mechanics
MO	Molecular Orbital
NDDO	Neglect of Diatomic Differential Overlap
NMR	Nuclear Magnetic Resonance
PAM	Point Accepted Mutations
PB	Penicillin-Binding
PBP	Penicillin-Binding Domain
PDB	Protein Data Bank
PM3	Parametric Method 3
PPP	Pariser-Parr-Pople
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid
SCF	Self-Consistent Field
SCOP	Structural Classification Of Proteins
SDS-PAGE	Sodiumdodecylsulfate Polyacrylamide Gel Electrophoresis
STO	Slater-Type Orbital
UDP	Uridine 5'-Pyrophosphate

## Appendix B Amino acids

Structures, three-letter codes and one-letter codes of the natural amino acids



## **Appendix C    Hardware**

The calculations presented in this thesis were performed on the following computers:

*At the Institute for Pharmaceutical Chemistry at the Heinrich-Heine Universität Düsseldorf:*

- SiliconGraphics Workstations
  - IRIS INDIGO<sup>2</sup> R10000 Solid Impact
  - IRIS INDY R5000
  - IRIS O2 R5000

*At the Rechenzentrum at the Heinrich-Heine Universität Düsseldorf:*

- SiliconGraphics Origin2000 with 32 R10000 processors

## Appendix D    Software

The following software was used :

<b>AutoDock2.4</b>	Docking program for docking flexible molecules into rigid proteins
<b>EXCEL</b>	Table calculation program for performing linear and non-linear regression analyses
<b>InsightII</b>	Molecular modelling software package from MSI/BIOSYM
<b>SPARTAN</b>	Software package for performing quantum chemical, semiempirical and molecular mechanical calculations from Wavefunctions
<b>SYBYL</b>	Molecular modelling software package from Tripos



## Appendix E Additional CVFF parameters

### *Bond angle parameters*

*Additional parameters for simulation of  $\beta$ -lactam antibiotics in the CVFF [189]*

#quadratic\_angle cvff  
> E = K2\* (Theta – Theta0)^2

I	J	K	Theta0	K2 [kcal.mol <sup>-1</sup> .rad <sup>-1</sup> ]
h	cB1	cB2	115.8111	48.9545
h	cB1	s	109.7109	49.3226
h	cB1	n	112.8128	49.2224
h	cB2	cB1	110.6913	49.1124
h	cB2	c'''	111.1866	48.8221
h	cB2	n	109.1610	49.4324
cB2	cB1	s	116.9587	58.2402
cB1	cB2	n	118.6774	123.1381
c'''	cB2	n	120.4056	49.1902
cB2	c'''	o''	135.7523	51.0141
s	cB1	n	111.9880	71.1461
c''	n	cB1	94.8484	71.8291
n	cB1	cB2	88.0173	75.8733
cB1	cB2	c'''	84.7865	113.5060
cB2	c'''	n	91.3899	78.9951
o''	c'''	n	132.8034	51.0192

*Additional parameters for simulation of the thiazole-ring system in the CVFF [188]*

I	J	K	Theta0	K2
sth	cth	h	119.9000	30.2723
h	cth	c5	124.0000	30.0000
cth	nth	c5	110.6000	60.9900
nth	cth	sth	115.3530	75.7400
nth	cth	n3	123.8990	51.3200
nth	c5	c	120.4100	51.2000
cth	sth	cth	88.2200	149.1700
nth	c5	cth	116.8200	53.6700
sth	cth	n3	121.0300	52.1900
cth	n3	hn	120.3450	49.0500
sth	cth	c	125.1150	50.4300
sth	cth	c5	108.9980	85.7000
c5	cth	c	125.8900	52.7100
cth	c5	c	122.7700	52.4400
cth	c	h	110.9060	49.0700
cth	c	c	109.4720	111.6100

**Torsion angle parameters**

*Additional torsion angle parameters for simulation of the thiazole-ring system in the CVFF [188]*

```
#torsion_1 cvff
> E = Kphi * [ 1 + cos(n*Phi - Phi0) ]
```

I	J	K	L	Kphi	n [kcal.mol <sup>-1</sup> ]	Phi0
c	c'	n	c=	3.2000	2	180.0000
o'	c'	n	c=	3.8000	2	180.0000
*	cth	nth	*	11.0218	2	180.0000
*	cth	sth	*	2.1114	2	180.0000
*	n3	cth	*	3.9060	2	180.0000
*	c	cth	*	0.9514	2	180.0000
*	nth	cp	*	4.0000	2	180.0000
*	cth	cp	*	6.0000	2	180.0000

**Out-of-plane angle parameters**

*Additional out-of-plane angle parameters for simulation of the thiazole-ring system in the CVFF [189]*

```
#out_of_plane cvff
> E = Kchi * [ 1 + cos(n*Chi - Chi0) ]
```

I	J	K	L	Kchi	n [kcal.mol <sup>-1</sup> ]	Chi0
np	cs	n3	sp	0.3700	2	180.0000
c	cs	cp	sp	2.9998	2	180.0000
c	cp	cs	np	0.3700	2	180.0000
hn	n3	hn	cs	0.0000	2	180.0000
hn	np	cp	np	0.3700	2	180.0000
np	cp	np	n3	0.3700	2	180.0000

## Appendix F BCL macros

The following macros have been written in the **biosym** command language (**BCL**).

### Auto\_replace.bcl

```
# Provided that a file containing the mutant protein sequence in three-  
# character code exists, this macro enables quick substitution of the  
# complete protein into a poly-alanine-glycine-proline mutant. The complete  
backbone and C resulting mutant protein protein
```

```
Define_Macro Auto_replace \  
  Ident Molecule_name \  
  Ident      pbp2x_seq  
  
  Int  N  
  Int  Row  
  Ident ID  
  Ident      Cell  
  Lstring      contents  
  Lstring contents  
  Set_Param_Pick Molecule_name MOLECULE_NAME  
  
  N = 76  
  Row = 1  
  Biopolymer  
  While ($N < 751)  
  
    ID = $Molecule_name // ":" // $N  
  
    Cell = $pbp2x_seq // ":" // $Row // ".1"  
    contents = {Cell_Get_String $Cell}  
    print GLY $N $contents  
    Replace Residue $ID $contents L  
  
    Potentials Fix -Print_Potentials Fix -Print_Part_Chargs \  
      Fix -Print_Form_Chargs $Molecule_name  
    N = $N + 1  
    Row = $Row + 1  
  End  
  
End_Macro  
Add_To_Pulldown Auto_replace User
```

**Arc\_to\_pdb.bcl**

```
# This macro quickly converts the various frames of a molecular dynamics
# run into pdb files for analysis with PROCHECK. In addition to removing
# the hydrogen atoms, the macro, which is specifically written for PBP2x,
# deletes both the N-terminal and C-terminal domains.
```

```
Define_Macro Arc_to_pdb \
  Ident Molecule_Name \
  Int   First_frame \
  Int   Last_frame
  Int   fnum
  Int   optnum
  Int   N
  Ident Res_Spec
  Sstring   Arc_PDB

  Set_Param_Pick Molecule_Name MOLECULE_NAME
  Arc_Name = "res_free.arc"
  optnum = $Last_frame + 1
  fnum = $Last_frame
  N = 1

  While ($fnum >= $First_frame)

      Mol_ID = $Molecule_Name//$optnum

      Object_ID = $Molecule_Name//$N

      Get Molecule Archive Frame $optnum $Arc_Name $Object_ID
      -Reference_Object

      Res_Spec = $Object_ID//":266-616"

      Display Molecule Only Atoms Specified $Res_Spec

      Display Molecule Off Atoms Hydrogens $Res_Spec

      Arc_PDB = "arc_"//$fnum//"_opt"

      Put Molecule PDB $Object_ID $Arc_PDB -Transformed Displayed -
Insight_Style

      fnum = $fnum - 1
      optnum = $optnum + 1
      N = $N + 1

  End

End_Macro
Add_To_Pulldown Arc_to_pdb User
```

**List\_CA\_angles.bcl**

```
# This macro lists the angles and dihedrals between the Cα atoms of a
# protein and lists them in a file trace_angles.tbl. Upon activation the
# macro asks for the first and last residues of the amino acid sequence for
# which these angle values have to be listed.
```

```
Define_Macro List_CA_angles \
  Ident Molecule_Name \
  Int Starting_Residue \
  Int Ending_Residue
  Float torsion_value
  Float angle_value

  presidue = $Starting_Residue - 1
  cresidue = $Starting_Residue
  nresidue = $Starting_Residue + 1
  nnresidue = $Starting_Residue + 2
  ang = trace_angles.tbl

  Set_Param_Pick Molecule_Name MOLECULE_NAME

  If ( ($Starting_Residue <= 1) || ($Ending_Residue <= 1) )
    Print "Residue number must be greater than 1"
    Return
  End

  If ( ($Starting_Residue > $Ending_Residue) || \
    ( ($Ending_Residue - $Starting_Residue) < 3) )
    Print "Ending_Residue number should be greater than"
    Print "Starting_Residue number by at least three."
    Return
  End

  # Remove trace_angles file, if present
  Bcl_Unix ("/bin/rm -f " // $ang)

  While ($nresidue < ($Ending_Residue + 3))
    CA_presidue = $Molecule_Name//":"//$presidue//":CA"
    CA_cresidue = $Molecule_Name//":"//$cresidue//":CA"
    CA_nresidue = $Molecule_Name//":"//$nresidue//":CA"
    CA_nnresidue = $Molecule_Name//":"//$nnresidue//":CA"

    $torsion_value = \
      {Dihedral -monitor $CA_nnresidue $CA_nresidue \
        $CA_cresidue $CA_presidue}
    $angle_value = \
      {Angle -monitor $CA_nresidue $CA_cresidue \
        $CA_presidue}

    Write $ang "%5d %11.6f %11.6f \n" $cresidue $torsion_value \
      $angle_value

    presidue = $cresidue
    cresidue = $nresidue
    nresidue = $nnresidue
    nnresidue = $nnresidue + 1
  End

  Close $ang

End_Macro
Add_To_Pulldown List_CA_angles User
```

**Fold\_trace.bcl**

# This macro folds the trace of any 3D amino acid chain according to the angle values present in a protein crystal structure.

```

Define_Macro Fold_trace \
  Ident Molecule_name \
  Ident trace_angles \
  Int Starting_Residue \
  Int Ending_Residue

  presidue      = $Starting_Residue - 1
  cresidue      = $Starting_Residue
  nresidue      = $Starting_Residue + 1
  nnresidue     = $Starting_Residue + 2

  Set_Param_Pick Molecule_Name MOLECULE_NAME

  If ( ($Starting_Residue <= 1) || ($Ending_Residue <= 1) )
    Print "Residue number must be greater than 1"
    Return
  End
  If ( ($Starting_Residue > $Ending_Residue) || \
    ( ($Ending_Residue - $Starting_Residue) < 3) )
    Print "Ending_Residue number should be greater than"
    Print "Starting_Residue number by at least three."
    Return
  End
  Row = 1
  Builder
  While ($nresidue < ($Ending_Residue + 3))
    CA_presidue = $Molecule_Name//":"//$presidue//":CA"
    CA_cresidue = $Molecule_Name//":"//$cresidue//":CA"
    CA_nresidue = $Molecule_Name//":"//$nresidue//":CA"
    CA_nnresidue = $Molecule_Name//":"//$nnresidue//":CA"

    Cell = $trace_angles // ":" // $Row // ".2"
          contents = {Cell_Get_String $Cell}

    Print $contents $CA_nnresidue $CA_nresidue $CA_cresidue \
          $CA_presidue

    Geometry Dihedral $contents $CA_nnresidue \
          $CA_nresidue $CA_cresidue $CA_presidue

    Cell = $trace_angles // ":" // $row // ".3"
          contents2 = {Cell_Get_String $Cell}

    Print $contents2 $CA_nresidue $CA_cresidue $CA_presidue

    Geometry Angle $contents2 $CA_nresidue \
          $CA_cresidue $CA_presidue

    presidue    = $cresidue
    cresidue    = $nresidue
    nresidue    = $nnresidue
    nnresidue   = $nnresidue + 1
  Row = $Row + 1
  End
End_Macro
Add_To_Pulldown Fold_trace User

```

**List\_CA\_distances.bcl**

```
# This macro measures the distances between the Cα atoms of a protein and
# lists them in a file ca_distances.tbl. Upon activation the macro asks for
# the first and last residues of the amino acid sequence for which these
# distances have to be listed.
```

```
Define_Macro List_CA_distances \
  Ident Molecule_Name \
  Int Starting_Residue \
  Int Ending_Residue
  Float distance_value

  cresidue = $Starting_Residue
  nresidue = $Starting_Residue + 1
  dis = ca_distances.tbl

  Set_Param_Pick Molecule_Name MOLECULE_NAME

  # check input parameters
  If ( ($Starting_Residue <= 1) || ($Ending_Residue <= 1) )
    Print "Residue number must be greater than 1"
    Return
  End

  If ( ($Starting_Residue > $Ending_Residue) || \
    ( ($Ending_Residue - $Starting_Residue) < 1) )
    Print "Ending_Residue number should be greater than"
    Print "Starting_Residue number by at least one."
    Return
  End

  # Remove ca_distances file, if present
  Bcl_Unix ("/bin/rm -f " // $dis)

  While ($cresidue < ($Ending_Residue + 1))

    CA_cresidue = $Molecule_Name//":"://$cresidue//:CA"
    CA_nresidue = $Molecule_Name//":"://$nresidue//:CA"

    $distance_value = \
      {Distance -monitor $CA_nresidue $CA_cresidue}

    Write $dis "%5d %11.6f \n" $cresidue $distance_value

    cresidue = $nresidue
    nresidue = $nresidue + 1

  End

  Close $dis

End_Macro
Add_To_Pulldown List_CA_distances User
```

**Modify\_CA\_distances.bcl**

```
# This macro modifies the distance between two Cα atoms according to a value
# read from a file by subsequently breaking the amide bond, adjusting the
# distance between the Cα atoms, and finally connecting the amide nitrogen
# and carbonyl oxygen atoms again. The overall structure of the protein is
# maintained, because the amino acids after the amide bond to be adjusted
# are translated by the same vector.
```

```
Define_Macro Modify_CA_distances \
  Ident Molecule_name \
  Ident ca_distances \
  Int Row \
  Int Col \
  Int Starting_Residue \
  Int Ending_Residue

  cresidue = $Starting_Residue
  nresidue = $Starting_Residue + 1

  Set_Param_Pick Molecule_Name MOLECULE_NAME

  If ( ($Starting_Residue <= 1) || ($Ending_Residue <= 1) )

    Print "Residue number must be greater than 1"
    Return
  End

  If ( ($Starting_Residue > $Ending_Residue) || \
    ( ($Ending_Residue - $Starting_Residue) < 1) )

    Print "Ending_Residue number should be greater than"
    Print "Starting_Residue number by at least one."
    Return
  End

  Row = 1
  Biopolymer
  While ($nresidue < ($Ending_Residue + 1))

    N_nresidue = $Molecule_Name//":"://"$nresidue//":N"
    C_cresidue = $Molecule_Name//":"://"$cresidue//":C"
    CA_cresidue = $Molecule_Name//":"://"$cresidue//":CA"
    CA_nresidue = $Molecule_Name//":"://"$nresidue//":CA"

    Bond Break $N_nresidue $C_cresidue

    Cell = $ca_distances // ":" // $Row // ".2"
    contents = {Cell_Get_String $Cell}

    Print $contents $CA_nresidue $CA_cresidue

    Geometry Distance $contents $CA_nresidue $CA_cresidue

    Bond Create -Fragment_Window "Partial Double" $N_nresidue \
      $C_cresidue -Assimilate

    cresidue = $nresidue
    nresidue = $nresidue + 1
    Row = $Row + 1
  End

End_Macro
Add_To_Pulldown Modify_CA_distances User
```



**RMS\_his\_trace.bcl**

```
# This macro calculates the RMSD between the starting conformation and all
# the frames generated during a molecular dynamics simulation, and writes
# these values into a table. Molecule and history files must be read before
# the macro is activated. The atoms that are to be fitted are defined in
# line 15 (only the C-alpha atoms of the transpeptidase domain of PBP2x
# have been specified here).
Define_Macro RMS_his_trace \
Lstring Molecule_name \
Int frames

Int index
float rmsd

Set_Param_Pick Molecule_name MOLECULE_NAME
Write ($Molecule_name//"_RMS.tab") "%s" "Frame" "\t"
Write ($Molecule_name//"_RMS.tab") "%s" "RMSD" "\n"
Copy Object -To_Clipboard -Displace $Molecule_name Templ
Foreach $index From 1 to $frames
  Conformation Trajectory Frame $index
  print $index
  print $Molecule_name
  print Templ
  Superimpose -End_Definition Trace "Label Mode" $Molecule_name:266-616
  Templ:266-616
  rmsd = {Superimpose End_Definition}
  Write ($Molecule_name//"_RMS.tab") "%d" $index "\t"
  Write ($Molecule_name//"_RMS.tab") "%f" $rmsd "\n"
End
Delete Object Templ
Conformation Trajectory Original

End_Macro

Add_To_Pulldown RMS_his_trace User
```

**RMS\_his\_ligand.bcl**

```
# This macro calculates the RMSD of the ligand / substrate / inhibitor
# (indicated here by CTXCOM_4) between the starting conformation and all
# the frames generated during a molecular dynamics simulation, and writes
# these values into a table.
# Molecule and history files must be read before the macro is activated.
Define_Macro RMS_his_ligand \
Lstring Molecule_name \
Lstring Template      \
Int      frames
Ident   Res_Spec1
Ident   Res_Spec2

Int index
float rmsd

Set_Param_Pick Molecule_name MOLECULE_NAME
Write ($Molecule_name//"_RMS.tab") "%s" "Frame" "\t"
Write ($Molecule_name//"_RMS.tab") "%s" "RMSD" "\n"
Foreach $index From 1 to $frames
  Conformation Trajectory Frame $index
  Copy Object -To_Clipboard -Displace $Molecule_name Config
  print $index
  Res_Spec1 = $Template
  Res_Spec2 = CTXCOM_4
  Superimpose -End_Definition Heavy "Label Mode" $Res_Spec2 $Res_Spec1
  rmsd = {Superimpose End_Definition}
  Write ($Molecule_name//"_RMS.tab") "%d" $index "\t"
  Write ($Molecule_name//"_RMS.tab") "%f" $rmsd "\n"
  Delete Object Config
End

End_Macro

Add_To_Pulldown RMS_lig User
```

**Interaction.bcl**

```
# This macro calculates the coulomb, van der Waals and total interaction
# enegies between the ligand / substrate / inhibitor and the protein of
# each frame generated during a molecular dynamics simulation, and writes
# these values into tables.
```

```
Define_Macro Interaction \  
Lstring Molecule_name \  
Lstring Molecule_1 \  
Lstring Molecule_2 \  
Int frames  
Float coul_value  
Float vdw_value  
Float tot_value  
Float derv_value  
Ident Res_Spec1  
Ident Res_Spec2  
Int index  
inter_vdw = E_VdW.tbl  
inter_coul = E_Coulomb.tbl  
inter_tot = E_Total.tbl  
Set_Param_Pick Molecule_name MOLECULE_NAME  
Foreach $index From 1 to $frames  
  Conformation Trajectory Frame $index  
  print $index  
  Docking  
  $coul_value = \  
    {Intermolecular -Monitor $Molecule_1 $Molecule_2 -Grid 100 \  
    -Derivatives -Van_der_waals Coulomb}  
  Write $inter_coul "%5d %11.6f \n" $index $coul_value  
  $vdw_value = \  
    {Intermolecular -Monitor $Molecule_1 $Molecule_2 -Grid 100 \  
    -Derivatives Van_der_waals -Coulomb}  
  Write $inter_vdw "%5d %11.6f \n" $index $vdw_value  
  $tot_value = \  
    {Intermolecular -Monitor $Molecule_1 $Molecule_2 -Grid 100 \  
    -Derivatives Van_der_waals Coulomb}  
  Write $inter_tot "%5d %11.6f \n" $index $tot_value  
  Analysis  
End  
Close $inter_vdw  
Close $inter_coul  
End_Macro  
Add_To_Pulldown Interaction User
```

## Appendix G    Awk scripts

### Extract\_E\_values.awk

```
# This awk script extracts interaction energy values from a .ref file
# generated during a molecular dynamics run.
awk '$1 == 2 && $2 == 3 {print $7}' autocom.ref > list_Coul.txt
```

### Make\_E\_table.awk

```
# This awk script puts the energy values extracted in Extract_E_values.awk
# into a table.
{i++}
{printf "%3s\t %10s\n", i-1, $1}
```

## Appendix H Analysis of molecular dynamics simulations

The tables shown here represent an analysis of the last six MD simulations in the generation of a final model of the transpeptidase domain of PBP2x. The following criteria were used:

```
Dynamics:      * T = 310 K
                * D = 4.0*r
                * cutoff = 13.5 - 15.0 A
                * run:      1) 20 ps initialisatie
                           2) 200 ps run
                * template: otto.tem

Optimisation:  * D = 4.0*r
                * cutoff = 13.5 - 15.0 A
                * template force (0 kcal/A)
                * 500 iterations steepest descent
                * conjugate gradients tot dE < 1.0 kcal/A
                * template: otto.tem
```

### Explanation:

sim tim	=	simulation time
E total	=	total energy of the optimised frame
%mfr	=	percentage of the amino acids in the most favoured regions
%dis	=	percentage of the amino acids occurring in the disallowed regions
cis	=	amino acids in the cis conformation
dis mcb	=	number of disallowed main chain bonds
dis mca	=	number of disallowed main chain angles
dis plg	=	number of disallowed planar groups
res_06_12	=	MD run during which the C-alpha atoms have been restrained with a force constant of 0.6 to 1.2 kcal/mol/Å.

The structures that have been selected as starting conformation for the next MD run are shown in bold.

**Res\_10\_20**

sim tim	E total	%mfr	% dis	cis	dis mcb	dis mca	dis plg
40	2387.321080	80.6	0.7	P308	7	0	3
60	2369.117342	81.9	0.7	P308	7	0	2
80	2380.842738	82.2	0.7	P308	9	0	4
100	2365.935286	79.6	0.7	P308	7	0	1
120	2377.376914	80.9	0.7	P308	9	0	2
140	2373.451340	80.9	0.7	P308	10	0	1
160	2367.925374	79.6	0.7	P308	8	0	3
180	2378.549966	78.0	0.7	P308	6	0	4
200	2332.613413	78.9	0.7	P308	6	0	2
<b>220</b>	<b>2372.866523</b>	<b>77.0</b>	<b>0.7</b>	<b>P308</b>	<b>5</b>	<b>0</b>	<b>4</b>

**res\_08\_16**

sim tim	E total	RMS PBD	%mfr	% dis	cis	dis mcb	dis mca	dis plg
40	2338.656719	0.49062	78.6	0.7	P308	4	0	1
60	2332.166061	0.46668	79.3	0.7	P308	3	0	3
80	2336.956220	0.44456	78.9	0.7	P308	5	0	1
100	2325.554648	0.44397	78.9	0.7	P308	5	0	3
120	2348.320409	0.42759	80.9	0.7	P308	5	0	1
140	2352.627963	0.41515	78.6	0.7	P308	7	0	3
160	2342.604762	0.39347	78.3	0.7	P308	5	0	4
180	2341.046470	0.43930	79.3	0.7	P308	3	0	3
<b>200</b>	<b>2354.745074</b>	<b>0.42578</b>	<b>80.9</b>	<b>0.7</b>	<b>P308</b>	<b>6</b>	<b>0</b>	<b>3</b>
220	2344.424524	0.40834	79.9	0.7	P308	5	0	4

**res\_06\_12**

sim tim	E total	RMS PBD	%mfr	% dis	cis	dis mcb	dis mca	dis plg
40	2319.550295	0.42137	78.9	0.7	P308	6	0	1
60	2317.268536	0.42278	79.9	0.7	P308	5	0	2
80	2328.758408	0.39608	80.6	0.7	P308	6	0	1
100	2332.561843	0.37755	78.3	0.7	P308	6	0	2
120	2345.374065	0.38696	79.3	0.7	P308	5	0	3
140	2337.647026	0.47848	80.3	0.7	P308	6	0	1
160	2344.803250	0.36967	80.9	0.7	P308	6	0	4
180	2356.898876	0.49198	80.6	0.7	P308	8	0	3
200	2329.366680	0.48031	79.9	0.7	P308	9	0	3
<b>220</b>	<b>2331.354050</b>	<b>0.40855</b>	<b>82.6</b>	<b>0.7</b>	<b>P308</b>	<b>6</b>	<b>0</b>	<b>2</b>

**res\_04\_08**

sim	E	RMS	%mfr	%	cis	dis	dis	dis
tim	total	PBD		dis		mcb	mca	plg
---	-----	-----	----	---	----	---	---	---
40	2315.379938	0.35769	80.6	0.7	P308	7	0	2
60	2303.623486	0.39650	80.6	0.7	P308	8	0	1
80	2319.156489	0.37252	79.3	0.7	P308	10	0	2
100	2321.729290	0.39468	79.9	0.7	P308	6	0	0
120	2320.717191	0.45319	78.0	0.7	P308	9	0	2
140	2308.338545	0.34769	80.9	0.7	P308	8	0	2
160	2318.308629	0.39725	80.3	0.7	P308	7	0	5
180	2312.761465	0.41557	81.9	0.7	P308	10	0	1
200	2320.331403	0.45058	81.6	0.7	P308	9	0	2
<b>220</b>	<b>2305.075077</b>	<b>0.44088</b>	<b>80.3</b>	<b>0.7</b>	<b>P308</b>	<b>6</b>	<b>0</b>	<b>3</b>

**res\_02\_04**

sim	E	RMS	%mfr	%	cis	dis	dis	dis
tim	total	PBD		dis		mcb	mca	plg
---	-----	-----	----	---	----	---	---	---
40	2306.351049	0.46697	77.0	1.0	P308	7	0	2
60	2301.041140	0.46834	78.3	1.0	P308	7	0	0
80	2306.783954	0.42748	78.6	1.0	P308	8	0	2
100	2312.608645	0.43504	78.9	1.0	P308	4	0	2
120	2322.141162	0.40386	78.9	1.0	P308	4	0	1
140	2310.048842	0.47057	78.6	1.0	P308	5	0	1
160	2305.097400	0.44121	79.6	1.0	P308	5	0	1
180	2312.703115	0.40608	78.3	1.0	P308	3	0	0
200	2298.501712	0.43091	77.6	1.0	P308	8	0	1
<b>220</b>	<b>2290.022097</b>	<b>0.43642</b>	<b>79.3</b>	<b>1.0</b>	<b>P308</b>	<b>8</b>	<b>0</b>	<b>0</b>

**res\_00**

sim	E	RMS	%mfr	%	cis	dis	dis	dis
tim	total	PBD		dis		mcb	mca	plg
---	-----	-----	----	---	----	---	---	---
40	2300.941204	0.47349	76.0	1.0	P308	6	0	3
60	2305.250057	0.39623	80.9	1.3	P308	10	0	0
80	2302.358367	0.38163	78.0	1.3	P308	9	0	1
100	2300.787910	0.41951	77.6	1.0	P308	7	0	1
120	2307.574989	0.45171	78.9	1.0	P308	9	2	1
140	2301.467248	0.39173	78.6	1.0	P308	7	0	0
160	2312.018429	0.40852	80.3	1.3	P308	11	1	2
<b>180</b>	<b>2282.762082</b>	<b>0.44307</b>	<b>79.3</b>	<b>1.0</b>	<b>P308</b>	<b>7</b>	<b>0</b>	<b>1</b>
200	2304.654408	0.48827	79.6	1.0	P308	8	0	0
220	2306.693631	0.48584	78.0	1.0	P308	10	0	2

## Appendix I      Docking scripts for AutoDock 2.4

The following scripts are used for docking with AutoDock 2.4. The input files are represented by `receptor.mol2` and `ligand.mol2`.

### script\_prepare\_mol

```
# AutoDock_2.4
# Generates ligand and receptor files (.pdbq format)
# torsion angle definition

autotors -m ligand.mol2 ligand.pdbq

# verification of charges

check-qs ligand
jot ligand.err
jot ligand.pdbq

# receptor file conversion
# charge verification

rem-lp receptor.mol2
mol2topdbq receptor.mol2-lp > receptor.pdbq
check-qs receptor
jot receptor.err
```

### script\_grid

```
# generates grid parameter files

prepare-gpf+dpf receptor ligand
jot ligand.dpf
autogrid -p ligand.gpf -l ligand.glg &
```

### script\_clamp

```
# filters grid fields

clamp receptor_C.map > receptor_C.map.new
clamp receptor_H.map > receptor_H.map.new
clamp receptor_N.map > receptor_N.map.new
clamp receptor_e.map > receptor_e.map.new

mv receptor_C.map.new receptor_C.map
mv receptor_H.map.new receptor_H.map
mv receptor_N.map.new receptor_N.map
mv receptor_e.map.new receptor_e.map
```



**script\_dock**

```
# actual docking
autodock -k -p ligand.dpf -l ligand.dlg &
```

**tool**

```
# generates lig files (no connectivity!!)
cat ligand.dlg | nawk -f prog | nawk -f program1
```

**program1**

```
BEGIN {s=1;file = "lig" s}
/REMARK/,/END/ {if ($1 == "END") {s++ ;print >> file ; print s; {file =
"lig" s}} else print >> file}
END {print file}
```

**script\_prep\_lig**

```
# preparation of more ligands for the same receptor
# torsion angle definition

autotors -m ligand.mol2 ligand.pdbq

# charge verification

check-qs ligand
jot ligand.err
jot ligand.pdbq
```

**script\_newjob**

```
# new docking with same receptor

prepare-gpf+dpf receptor ligand
jot ligand.dpf
echo 'autodock -k -p ligand.dpf -l ligand.dlg &'
```

## Appendix J Input files for DISCOVER

### Example of an input file for performing simulated annealing

```
! INPUT FILE FOR DISCOVER GENERATED BY PAUL
!
!
overlap = 0.01
cutoff = 13.500000
cutdis = 15.000000
swtdis = 1.5
DEMAX = 10000
!
begin simulation
* template file = "pmd_acts_mclus.tem"
* add-automatic bond torsion valence out-of-plane
  reduce
!
set dielectric = 3.500000*r
!
HBDIST = 3.5
HBANGL = 130.5
!
!
Fixed atom list generation
* add all
* molecule 1 residue ALAN 76 to ALA 265
* add all
* molecule 1 residue ALA 617 to ALAC 750
!
!
tethered atoms list generation
* add calpha
* molecule 1 residue SER 266 to ASP- 616
!
!
restrain calpha in molecule 1 residues SER 266 to ASP- 616
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
restrain heavy in molecule 1 residues SER 337 to ALA 347
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
restrain heavy in molecule 1 residue PHE 392
```

```
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
  restrain heavy in molecule 1 residues SER 395 to ASN 397
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
  restrain heavy in molecule 1 residue MET 400
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
  restrain heavy in molecule 1 residue MET 407
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
  restrain heavy in molecule 1 residue TRP 412
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
  restrain heavy in molecule 1 residue PHE 419
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
  restrain heavy in molecule 1 residues LYS+ 547 to GLY 549
* to corresponding atoms in template using 2 with a maximum force
* of 5 kcal/mole/A
!
!
  minimize
*   no cross terms
*   no morse
*   for 100000 iterations
*   using conjugate gradient
*   until the maximum derivative is less than 1.0000000 kcal/A
!
!
100  initialize dynamics at 610.0 for 20000 steps of 1 fs
* with no cross terms and no morse functions
* averages every 500 steps
* write history file every 1000 steps
!
  carnum = 1
200  resume dynamics at 610.0 for 10000 steps of 1 fs
* with no cross terms and no morse functions
  rms comparison
  archive as file number carnum
  carnum = carnum + 1
  if carnum .le. 6 then 200
!
```

```
    resume dynamics at 610.0 for 5000 steps of 1 fs
* with no cross terms and no morse functions
!
!
300  resume dynamics at 610.0 for 15000 steps of 1 fs
* with no cross terms and no morse functions
!
    temp = 610
    timtmp = 0.5
400  temp = temp-50
    resume dynamics at temp for 1000 steps of 1 fs
* with no cross terms and no morse functions
    if temp .gt. 310 then 400
!
500  resume dynamics at 310.0 for 15000 steps of 1 fs
* with no cross terms and no morse functions
!
    temp = 310
    timtmp = 0.5
600  temp = temp-50
    resume dynamics at temp for 1000 steps of 1 fs
* with no cross terms and no morse functions
    if temp .gt. 110 then 600
!
700  resume dynamics at 110.0 for 15000 steps of 1 fs
* with no cross terms and no morse functions
    archive as file number carnum
    rms comparison
    carnum = carnum + 1
    if carnum .le. 11 then 300
!
!
    carnum = 11
    cornum = 12
!
    restrain clear
!
800  retrieve as file number carnum
    template force for 500 cycles using
* steepest descent with a
* forcing constant of 0 kcal/A
* no cross terms
* no morse
!
    template force for 10000 cycles using
* conjugate gradients with a
* forcing constant of 0 kcal/A
* no cross terms
* no morse
```

```
* until the maximum derivative is less than 1.0 kcal/A
  archive as file number cornum
  rms comparison
!
900  print hbonds
* add all
* molecule 1 residue  SER  266 to  ASP- 616
  carnum = carnum - 1
  cornum = cornum + 1
  if carnum .ge. 1 then 800
!
!
  end
```

### **Example of an input file for performing a molecular dynamics simulation with both restraining and tethering forces**

*Although the atoms specified in the tethering list are NOT tethered during the simulation, they need to be specified this way, in order to define the template atoms.*

```
!   INPUT FILE FOR DISCOVER GENERATED BY PAUL
!
!   molecule 1 = structural water molecule
!   molecule 2 = PBP2x
!
  overlap = 0.01
  cutoff = 13.500000
  cutdis = 15.000000
  swtdis = 1.5
  DEMAX = 10000
!
  begin simulation
*   template file = "otto.tem"
*   add-automatic bond torsion valence out-of-plane
  reduce
!
  set dielectric = 4.000000*r
!
  HBDIST = 3.5
  HBANGL = 130.5
!
!
  Fixed atom list generation
```

```
* add all
* molecule 2 residue ALAN 76 to ALA 265
* add all
* molecule 2 residue ALA 617 to ALAC 750
!
!
tethered atoms list generation
* add calpha
* molecule 2 residue SER 266 to ASP- 616
!
!
!
restrain calpha in molecule 2 residues SER 266 to THR 309
* to corresponding atoms in template using 0.6 with a maximum force
* of 1.2 kcal/mole/A
!
restrain calpha in molecule 2 residues PHE 310 to TYR 333
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
restrain calpha in molecule 2 residues GLU- 334 to ASN 351
* to corresponding atoms in template using 0.6 with a maximum force
* of 1.2 kcal/mole/A
!
restrain calpha in molecule 2 residues THR 352 to ALA 369
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
restrain heavy in molecule 2 residues THR 370 to MET 386
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
restrain calpha in molecule 2 residues THR 387 to PHE 392
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
restrain calpha in molecule 2 residues ALA 393 to ARG+ 418
* to corresponding atoms in template using 0.6 with a maximum force
* of 1.2 kcal/mole/A
!
restrain calpha in molecule 2 residues PHE 419 to ASP- 440
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
restrain calpha in molecule 2 residues ASN 441 to MET 515
* to corresponding atoms in template using 0.6 with a maximum force
* of 1.2 kcal/mole/A
!
restrain calpha in molecule 2 residues VAL 516 to PRO 540
```

```
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
  restrain calpha in molecule 2 residues GLY  541 to SER  576
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
  restrain calpha in molecule 2 residues PRO  577 to ASP- 616
* to corresponding atoms in template using 0.6 with a maximum force
* of 1.2 kcal/mole/A
!
  restrain molecule 2 residue THR  338 atom HG1 and molecule 1
* residue WTR  1 atom O1 to each other
* using 1 kcal/mole/ang**2 with maximum 2
!
  restrain molecule 2 residue TYR  586 atom HH and molecule 1
* residue WTR  1 atom O1 to each other
* using 1 kcal/mole/ang**2 with maximum 2
!
  restrain molecule 2 residue SER  571 atom OG and molecule 1
* residue WTR  1 atom H12 to each other
* using 1 kcal/mole/ang**2 with maximum 2
!
  restrain molecule 2 residue PRO  335 atom O and molecule 1
* residue WTR  1 atom H11 to each other
* using 1 kcal/mole/ang**2 with maximum 2
!
  restrain side in molecule 2 residue LYS+ 547
* to corresponding atoms in template using 0.6 with a maximum force
* of 1.2 kcal/mole/A
!
  print restraints
!
!
!
100  minimize
*   no cross terms
*   no morse
*   for 10000 iterations
*   using conjugate gradient
*   until the maximum derivative is less than 1.0000000 kcal/A
!
!
!
200  initialize dynamics at 310.0 for 20000 steps of 1 fs
* with no cross terms and no morse functions
* averages every 500 steps
* write history file every 1000 steps
!
```

```
        carnum = 1
!
300  resume dynamics at 310.0 for 20000 steps of 1 fs
    * with no cross terms and no morse functions
    rms comparison
    archive as file number carnum
    carnum = carnum + 1
    if carnum .le. 10 then 300
!
!
!
!
    restrain
    * clear
!
    print restraints
!
!
    carnum = 10
    cornum = 11
!
400  retrieve as file number carnum
    template force for 500 cycles using
    * steepest descent with a
    * forcing constant of 0 kcal/A
    * no cross terms
    * no morse
!
500  template force for 10000 cycles using
    * conjugate gradients with a
    * forcing constant of 0 kcal/A
    * no cross terms
    * no morse
    * until the maximum derivative is less than 1.0 kcal/A
    archive as file number cornum
    rms comparison
!
600  print hbonds
    * add all
    * molecule 1 residue  WTR  1
    carnum = carnum - 1
    cornum = cornum + 1
    if carnum .ge. 1 then 400
!
!
!
    end
```



**Example of an input file for performing a simulated annealing molecular dynamics simulation of a complex with both restraining and tethering forces. Each picosecond the interaction energy is logged to a file (.ref)**

```
! INPUT FILE FOR DISCOVER GENERATED BY PAUL
!
! molecule 1 = structural water molecule behind B-sheet
! molecule 2 = PBP2x
! molecule 3 = structural water molecule in active site
! molecule 4 = benzylpenicillin
!
!
! overlap = 0.01
! cutoff = 33.000000
! cutdis = 34.500000
! swtdis = 1.5
! DEMAX = 10000
!
! begin simulation
! * template file = "otto.tem"
! * add-automatic bond torsion valence out-of-plane
! reduce
!
! set dielectric = 1.000000*r
!
! HBDIST = 3.5
! HBANGL = 130.5
!
!
! Fixed atom list generation
! * add all
! * molecule 2 residue ALAN 76 to ALA 265
! * add all
! * molecule 2 residue ALA 617 to ALAC 750
!
!
! tethered atoms list generation
! * add calpha
! * molecule 2 residue SER 266 to ASP- 616
!
!
!
! restrain calpha in molecule 2 residues PHE 310 to TYR 333
! * to corresponding atoms in template using 1 with a maximum force
! * of 2 kcal/mole/A
!
!
! restrain calpha in molecule 2 residues THR 352 to PHE 392
! * to corresponding atoms in template using 10 with a maximum force
```

```
* of 20 kcal/mole/A
!
  restrain calpha in molecule 2 residues PHE  419 to ASP- 440
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
  restrain calpha in molecule 2 residues VAL  516 to PRO  540
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
  restrain calpha in molecule 2 residues GLY  541 to PHE  570
* to corresponding atoms in template using 10 with a maximum force
* of 20 kcal/mole/A
!
  restrain calpha in molecule 2 residues GLN  590 to ASP- 616
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
!
  restrain side in molecule 2 residue SER  337
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
  restrain side in molecule 2 residue SER  395
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
  restrain molecule 2 residue PHE  392 atom O and molecule 3
* residue WTR  2 atom H12 to each other
* using 10 kcal/mole/ang**2 with maximum 20
!
  restrain molecule 2 residue ALA  393 atom O and molecule 3
* residue WTR  2 atom H11 to each other
* using 10 kcal/mole/ang**2 with maximum 20
!
  print restraints
!
!
!
  minimize
*   no cross terms
*   no morse
*   for 10000 iterations
*   using conjugate gradient
*   until the maximum derivative is less than 1.0000000 kcal/A
!
!
  log molecule-molecule interactions
!
```

```
    log energy/molecule
    * add molecule 1 to 4
!
!
!
100  initialize dynamics at 110.0 for 10000 steps of 1 fs
    * with no cross terms and no morse functions
    * averages every 500 steps
    * write history file every 1000 steps
!
!
    carnum = 1
!
!
    iter = 1
!
200  resume dynamics at 110.0 for 1000 steps of 1 fs
    * with no cross terms and no morse functions
    rms comparison
    log molecule-molecule interactions
    log energy/molecule
    * add molecule 1 to 4
    iter = iter + 1
    if iter .le. 10 then 200
!
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!
!
    iter = 1
    timtmp = 0.5
!
300  resume dynamics at 310.0 for 1000 steps of 1 fs
    * with no cross terms and no morse functions
    rms comparison
    log molecule-molecule interactions
    log energy/molecule
    * add molecule 1 to 4
    iter = iter + 1
    if iter .le. 20 then 300
!
!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!
!
    restrain
    * clear
!
    print restraints
```

```
!
    restrain calpha in molecule 2 residues PHE 310 to TYR 333
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
    restrain calpha in molecule 2 residues THR 352 to PHE 392
* to corresponding atoms in template using 10 with a maximum force
* of 20 kcal/mole/A
!
    restrain calpha in molecule 2 residues PHE 419 to ASP- 440
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
    restrain calpha in molecule 2 residues VAL 516 to PRO 540
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
    restrain calpha in molecule 2 residues GLY 541 to PHE 570
* to corresponding atoms in template using 10 with a maximum force
* of 20 kcal/mole/A
!
    restrain calpha in molecule 2 residues GLN 590 to ASP- 616
* to corresponding atoms in template using 1 with a maximum force
* of 2 kcal/mole/A
!
    restrain molecule 2 residue PHE 392 atom O and molecule 3
* residue WTR 2 atom H12 to each other
* using 1 kcal/mole/ang**2 with maximum 2
!
    restrain molecule 2 residue ALA 393 atom O and molecule 3
* residue WTR 2 atom H11 to each other
* using 1 kcal/mole/ang**2 with maximum 2
!
    print restraints
!
!
600  iter = 1
!
700  resume dynamics at 310.0 for 1000 steps of 1 fs
* with no cross terms and no morse functions
    rms comparison
    log molecule-molecule interactions
    log energy/molecule
* add molecule 1 to 4
    iter = iter + 1
    if iter .le. 20 then 700
    archive as file number carnum
    carnum = carnum + 1
    if carnum .le. 10 then 600
```

```
!  
!  
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!  
!  
!  
    restrain  
    * clear  
!  
    print restraints  
!  
!  
    carnum = 10  
    cornum = 11  
!  
800  retrieve as file number carnum  
    template force for 500 cycles using  
    * steepest descent with a  
    * forcing constant of 0 kcal/A  
    * no cross terms  
    * no morse  
!  
    template force for 10000 cycles using  
    * conjugate gradients with a  
    * forcing constant of 0 kcal/A  
    * no cross terms  
    * no morse  
    * until the maximum derivative is less than 1.0 kcal/A  
    rms comparison  
    archive as file number cornum  
    print molecule-molecule interactions  
    print energy/molecule  
    * add molecule 1 to 4  
!  
900  print hbonds  
    * add all  
    * molecule 4 residue  BPC  1  
    carnum = carnum - 1  
    cornum = cornum + 1  
    if carnum .ge. 1 then 800  
!  
!  
!  
!  
    end
```



---

An dieser Stelle möchte ich an alle, die mich während der Promotionszeit unterstützt, und somit entscheidend zum Gelingen dieser Arbeit beigetragen haben, nochmal ein ganz herzliches Dankeschön sagen. Besonders möchte ich mich bedanken bei:

allen Kolleginnen und Kollegen des Arbeitskreises für das ausgezeichnete Arbeitsklima, die stete Hilfsbereitschaft und die vielen gute Gespräche, aber insbesondere für die Hilfe, die mir wegen der Sprachprobleme am Anfang der Promotionszeit von allen angeboten wurde.

den Assistentinnen und Assistenten des vierten Semesters des Instituts für Pharmazie an der Heinrich-Heine-Universität Düsseldorf für die gute und kollegiale Zusammenarbeit. Besonders Herrn Dr. O. Grapenthin und Herrn Dr. R. Schulte gelten mein Dank für die Einweisung in das "Assistentensein" für Pharmazeuten. Herrn Dr. W. Sippl, Herrn C. Asche, Herrn F. Sunkel und Herrn Chönlau gelten mein Dank für die am Ende der Promotionszeit gezeigte Verlässlichkeit und Kollegialität.

Herrn C. Kirchhoff für die freundschaftliche Zusammenarbeit, die vielen konstruktiven Diskussionen und die aussergewöhnliche Hilfsbereitschaft bei den verschiedensten Computer-relatierten Problemen. Vor allem beim Posterdrucken waren seine Systemadministratorkenntnisse unfehlbar. Ein besonders großes Dankeschön auch für die Layout-Tips zu Beginn des Schreibens der Arbeit.

Herrn Priv.-Doz Dr. K.-J. Schleifer für seine stete Hilfe, und nicht zu vergessen seine praktischen Ratschläge bezüglich des Vaterseins.

Herrn Dr. W. Sippl für seine guten Ratschläge und vor allem seine Hilfe bei AutoDock.

Frau Dr. A. Schappach und Herrn B. Rupp für die zahllosen unterhaltsamen Diskussionen und die hilfreichen Anmerkungen.

Herrn H. Köbert und Herrn P. Sippel für die Betreuung der Institutsrechner, und Herrn P. Schreiber und Herrn R. Hartmann für die Betreuung des Silicon Graphics Origin2000 Parallelrechners am Rechenzentrum der Heinrich-Heine Universität Düsseldorf, dessen gelegentliches Abstürzen mich sonst gänzlich zur Verzweiflung gebracht hätte.

Herrn K. Lohmann, Herrn S. Stöber, Herrn G. und Frau S. Lang, Herrn C. Asche und Frau A. Schappach für die angenehmen frühmorgendlichen Schwimm- und Springstunden, wodurch ich bei der Arbeit einen einigermaßen kühlen Kopf habe behalten können.

meinem langjährigen Freund und guten Gesprächspartner R. Stok, der mit unseren, in regelmässigen Abständen stattgefundenen, manchmal sehr lange dauernden Schachspielen, für angenehme Abwechslung und Entspannung gesorgt hat.

meinen Eltern für die ständige Anteilnahme und das grosse Vertrauen, dass sie schon während meines Studiums, aber sicher auch während der Promotion in mich gestellt haben.

meinem Bruder Peter und seiner Freundin Agnes, die mich aus der Ferne ständig im Auge behalten, und mich mit ihren wertvollen Ratschlägen immer wieder stimuliert haben.

meiner Frau Julia, die mich in schwierigen Momenten immer aufgemuntert und durch ihre Fürsorge moralisch unterstützt hat. Ihre Geduld hat somit entscheidend zum Zustandekommen dieser Arbeit beigetragen.

meinem Sohn Laurens, dem ich aber erst in einigen Jahren deutlich machen werden kann, wie sehr er mich während der letzten Monate motiviert und durch seine fröhliche Art immer wieder aufgemuntert hat.



## Lebenslauf

<b>Name</b>		Paulus Albertus Maria van Hooft
<b>Geburtsdatum/~ort</b>		14. Juli 1969 in Utrecht
<b>Eltern</b>		Cornelis J. A. van Hooft Thérèse M. J. Maagdenberg-van Hooft
<b>Schulbildung</b>	08/75 - 07/81	Grundschule in Gouda
	08/81 - 06/88	Gymnasium in Gouda
	06/88	Abitur
<b>Studium</b>	09/88 - 06/95	Studium der Chemie an der Universität Utrecht
<b>Studienabschluss</b>	06/95	Chemiker
<b>Praktikum</b>	06/95 – 09/95	Praktikum bei N.V. Organon in Oss (Computational Medicinal Chemistry)
<b>Promotion</b>	09/96	Beginn der vorliegenden Dissertation am Institut für Pharmazeutische Chemie der Heinrich-Heine Universität Düsseldorf unter der Anleitung von Prof. Dr. H.-D. Höltje
<b>Berufserfahrung</b>	11/96 – 07/00	Wissenschaftlicher Mitarbeiter am Institut für Pharmazeutische Chemie an der Heinrich- Heine Universität Düsseldorf









