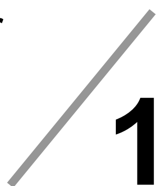
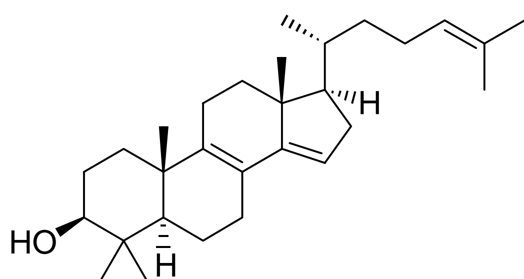


Chapter 1

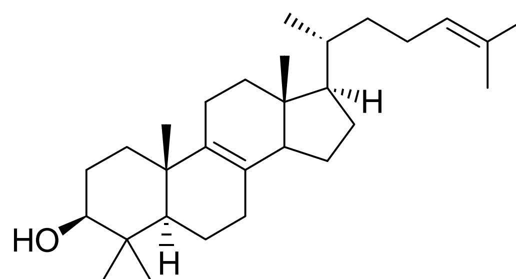
Introduction

Meiosis Activating Sterols

In 1995, Byskov and co-workers discovered two compounds that were found to be important for the development of mammalian germ cells. These steroid-like molecules are intermediates in the biosynthetic pathway of cholesterol, which are called sterols (Figure 1). They were found to exhibit a similar function: both induce immature reproductive cells to complete their development (Byskov *et al.*, 1995; Byskov *et al.*, 1998; Byskov *et al.*, 1999; Yding Andersen *et al.*, 1999). The molecules represent a class of compounds with novel pharmaceutical properties, designated Meiosis Activating Sterols (MAS) or, alternatively, Meiosis Inducing Substances (MIS). The names of the compounds are derived from their function and the location in the organisms they were isolated from. Compound **1**, 4,4-dimethyl-5α-cholest-8,14,24-triene-3β-ol, initiates the maturation of human oocytes in ovarian follicles and is called follicular fluid-MAS (FF-MAS). Compound **2**, 4,4-dimethyl-5α-cholest-8,24-diene-3β-ol, is involved in the production of spermatozoa in bull testis and is therefore called testicular-MAS (T-MAS).



4,4-dimethyl-5α-cholest-8,14,24-triene-3β-ol
(FF-MAS)



4,4-dimethyl-5α-cholest-8,24-diene-3β-ol
(T-MAS)

Figure 1. Structures of the meiosis activating compounds **1** (FF-MAS) and **2** (T-MAS).

The importance of the discovery of the two compounds lies in the possible applications of MAS-derived structures. Interference in the biological processes on which the MAS compounds operate can either promote or prevent fertility (Byskov *et al.*, 1999). Thus, MAS antagonists that block MAS action in either sex are attractive alternatives for the contraceptives used to date, since they are not likely to affect the normal hormonal balance. Furthermore, MAS agonists may prove valuable in treating problems related to reproductive cell maturation. Thus, sperm production in men could be increased and aberrant oocyte production in women could be remedied. The work described in this thesis is focused on the exploration of structural and physicochemical properties of MAS, in particular of FF-MAS and derived compounds. Implications of these

properties on the MAS activity are discussed and where possible, structure-activity relationships will be derived.

Meiosis

Meiosis is a reductive division process to produce reproductive cells with half the genetic material of the normal cell. During the meiotic process, the genetic content of a diploid precursor cell, with a single set of chromosomes and duplicate DNA, is reduced. The meiotic cell cycle, depicted in Figure 2, consists of four phases in which the cell grows, replicates DNA, segregates the chromosomes and divides. During the G1 phase the cell accumulates the enzymes needed to replicate DNA and other material that provides the necessary resources to complete the cycle. The S1 phase that follows is used to replicate the genetic material. The resulting cell, with two sets of chromosomes and therefore four DNA copies, then undergoes two successive reduction divisions called MI and MII. In mitosis, only one reduction division takes place, after which the resulting cells can enter a new division cycle starting with a second S phase. In meiosis, however, the development of the germ cells is continued and no S phase occurs between the reduction phases. In males, four haploid cells are produced, which all become functional sperm. In females, only one of the four haploids develops into a fully functional ovum, the rest of the genetic material is extruded as polar bodies (PB's). The development of the gametes can be arrested at several stages of the meiotic cycle, *i.e.* at the late G2/early MI phase, which is called prophase I and after the oocyte has developed into a fully functional egg cell.

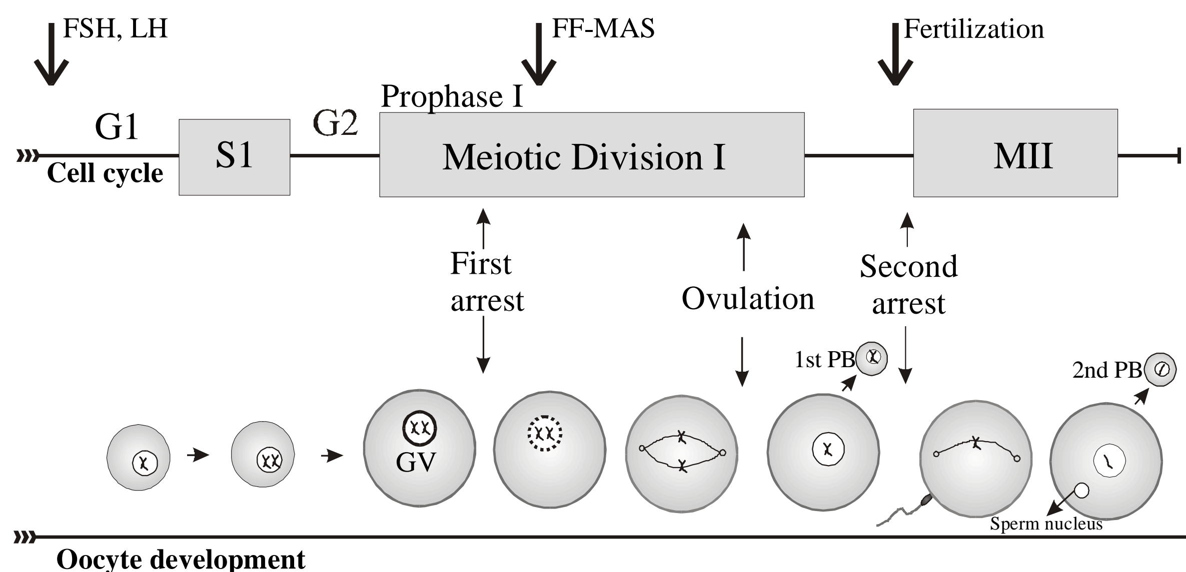


Figure 2. Schematic representation of the meiotic cell cycle.

In humans, the meiotic principle is the same for both sexes, but the timing is very different. Male germ cells are arrested in the G1 phase, and meiosis is not initiated until puberty. Thereafter, meiosis is continuous and uninterrupted. Female germ cells develop into primary oocytes during the first meiotic phase in fetal life. Their development is arrested a first time in the late prophase I stage, before or just after birth. This stage is characterized by a prominent nuclear membrane, called the germinal vesicle (GV), which can be seen through a microscope. The primary oocytes rest until puberty after which each in turn goes through the MI and MII phases. During this process, called GVBD, the germinal vesicle breaks down and the first polar body is extruded, which contains half of the DNA of the diploid mother cell. Thereafter, meiosis is arrested a second time and is resumed exclusively by entering of a sperm cell, after which the second PB is extruded.

Regulation of meiosis

The development of primary oocytes into egg cells takes place in follicles, which are the functional units of the ovary and ensure a unique environment. Primary oocytes arrested at the prophase stage are surrounded in succession by an antrum (*i.e.* fluid derived from serum), a layer of cumulus cells and a layer of mural granulosa cells. All of these cells are in communication through gap junctions, which facilitate transportation of hormones and lower molecular-weight compounds. Meiosis is initiated through a preovulatory surge of gonadotropins (Yding Andersen *et al.*, 1999), in particular follicle stimulating hormone (FSH) and luteinizing hormone (LH). The gonadotropins are also responsible for ovulation. Receptors for both hormones seem to be present on mural granulosa cells and cumulus cells exclusively, which makes them the primary target of these gonadotropins (Byskov *et al.*, 1997; Lu *et al.*, 2000). After the gonadotropin surge, the cyclic AMP level in the cumulus-oocyte complex increases (Eppig & Downs, 1988), whereas the cAMP level in the oocyte itself seems to decrease (Schultz, 1988). Meiosis does not occur when phosphodiesterase (PDE) enzymes, which are responsible for cAMP degradation, are inhibited in oocytes (Wiersma *et al.*, 1998). These findings indicate that a drop in intracellular cAMP concentration within the oocyte may be important for induction of meiosis. In fact, follicular fluid contains a PDE inhibitor, hypoxanthine, which is important in maintaining meiotic arrest (Törnell *et al.*, 1991; Hegele-Hartung *et al.*, 1999). In addition, it was found that phosphodiesterase 3 inhibitors blocked oocyte maturation in rodents, without affecting the normal cycle (Wiersma *et al.*, 1998).

Previously, it was believed that the resumption of meiosis was accomplished by withdrawing inhibitory substances from the oocytes, thus releasing it to continue meiosis spontaneously (Moor & Warnes, 1979; Dekel, 1995). However, after the gonadotropin surge, the hypoxanthine levels in mouse

follicular fluid levels drop only slightly, indicating that a stimulatory component may override the inhibitory effect of hypoxanthine. The discovery that MAS is able to induce maturation of the oocyte despite the presence of hypoxanthine (Törnell *et al.*, 1991; Hegele-Hartung *et al.*, 1999), indicates that such a mechanism is in fact present. The notion that MAS compounds are the endogenous activating stimulants of resumption of meiosis is confirmed by studies by Leonardsen *et al.* (2000a) and Lu *et al.* (2000). They showed that inhibition of different proteins involved in the synthesis of MAS prevents meiosis. In addition, it was found that 14α-demethylase, which demethylates lanosterol to give FF-MAS (see below), is important for meiosis (Yoshida *et al.*, 1996; Strömstedt *et al.*, 1998; Majdic *et al.*, 2000).

FF-MAS induces the resumption of meiosis in oocytes separated from their cumulus and granulosa cells, in contrast to FSH (Yding Andersen *et al.*, 1999). Thus, FF-MAS exerts in action on the oocyte directly. It was found that FSH is able to induce maturation only if the cumulus-enclosed oocyte complex is still intact just after the preovulatory surge of gonadotropins (Byskov *et al.*, 1997; Yding Andersen *et al.*, 1999). This suggests that FF-MAS is synthesized in the cumulus cells in response to elevated levels of FSH and transported to the oocyte, either through gap junctions, or in a paracrine way, *i.e.* by diffusion through the extracellular medium (the antrum). Communication between both cumulus cells and the oocyte appears to be necessary for the function of MAS. A receptor protein is believed to mediate the signal induced by MAS, but little is known about the receptor for MAS compounds. One possible candidate is the orphan nuclear receptor LXRα, which is activated five-fold by binding of FF-MAS (Janowski *et al.*, 1996; Ruan *et al.*, 1998). Oxycholesterol derivatives can also bind to the LXRα, but they are, in contrast to FF-MAS, not able to induce resumption of meiosis *in vitro* (Grøndahl *et al.*, 1998). Therefore, it does not seem likely that LXRα is the endogenous MAS receptor. A study on the influence of cholera toxin, a G-protein coupled receptor inhibitor, on the resumption of meiosis suggests that a G-protein-coupled receptor mechanism is responsible for the MAS mode of action (Grøndahl *et al.*, 2000). Recent studies suggest that the MAS receptor might be a plasma membrane-associated molecule (Færge *et al.*, 2001) involved in the cAMP-protein kinase A-dependent signal transduction pathway (Leonardsen *et al.*, 2000b).

Sterols: nomenclature and biosynthesis

Sterols, including the MAS compounds, are intermediates in the post-squalene biosynthesis of cholesterol. The discovery of the involvement of MAS compounds in the resumption of meiosis (Byskov *et al.*, 1995) was the first time that a specific function of a sterol compound, other than cholesterol, was found (Jackson *et al.*, 1997). Before discussion of the pathway involved in producing cholesterol, a brief description of the nomenclature and notations of sterols is

given. In Figure 3, the labeling scheme of the atoms of the androstane skeleton is given, as well as the names of the four rings of the sterol and steroid skeletons. The side above the plane of the paper of Figure 3 is called the β-side, the opposite side is called the α-side. Substituents occurring above the plane of the paper are denoted by a β, those occurring below the plane of the paper by an α following the atom number at which they occur. Double bonds are denoted by a Δ, with the atom numbers at which they occur as superscripts. If the double bond occurs between two consecutive numbered atoms, the higher atom number is not given explicitly. However, if this does not apply, the higher number is given between parentheses, directly after the lower atom number. If more than one double bond is present, the numbers that indicate their positions are separated by a comma. Thus, FF-MAS is a $\Delta^{8,14,24}$ and T-MAS a $\Delta^{8,24}$ structure.

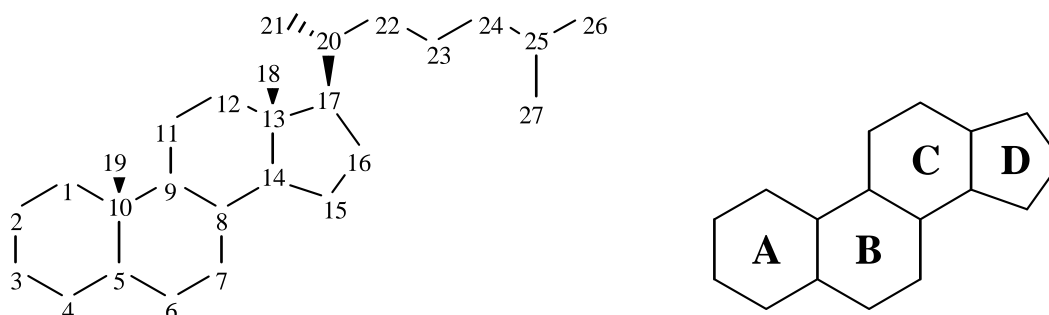


Figure 3. Numbering of the skeletal atoms in cholesterol and the MAS compounds (left) and the designation of the four skeleton rings (right). The β-side is above and the α-side is below the plane of the paper.

The post-squalene biosynthetic pathway of cholesterol is schematically represented in Figure 4 (Schroepfer Jr., 1982). Cyclization of squalene yields lanosterol, which is subsequently demethylated at C14 by cytochrome P450-14α-demethylase to give FF-MAS. Reduction of the Δ^{14} double bond by 14α-reductase yields T-MAS, which is then converted into zymosterol by elimination of two methyl groups at C4 through a number of oxidation and reduction steps. The Δ^{8} bond in zymosterol is converted to a Δ^{7} bond by a Δ^{8}-isomerase. Dehydrogenation of the single bond between C5 and C6 and reduction of the Δ^{7} double bond gives desmosterol. Reduction of the Δ^{24} double bond gives cholesterol. Alternatively, the Δ^{24} bond may be reduced before dehydrogenation of the Δ^{5} bond to give lathosterol. Dehydrogenation of the C5 to C6 single bond and subsequent reduction of the double bond at C7 gives cholesterol. The synthesis of FF-MAS and T-MAS follows the reverse direction of the biosynthetic pathway of cholesterol, starting from desmosterol ($\Delta^{5,7}$). During the reactions, several intermediate isomers can in theory be formed. It was observed that only a limited number could actually be obtained in appreciable yields. Apparently, the isomers have different stabilities.

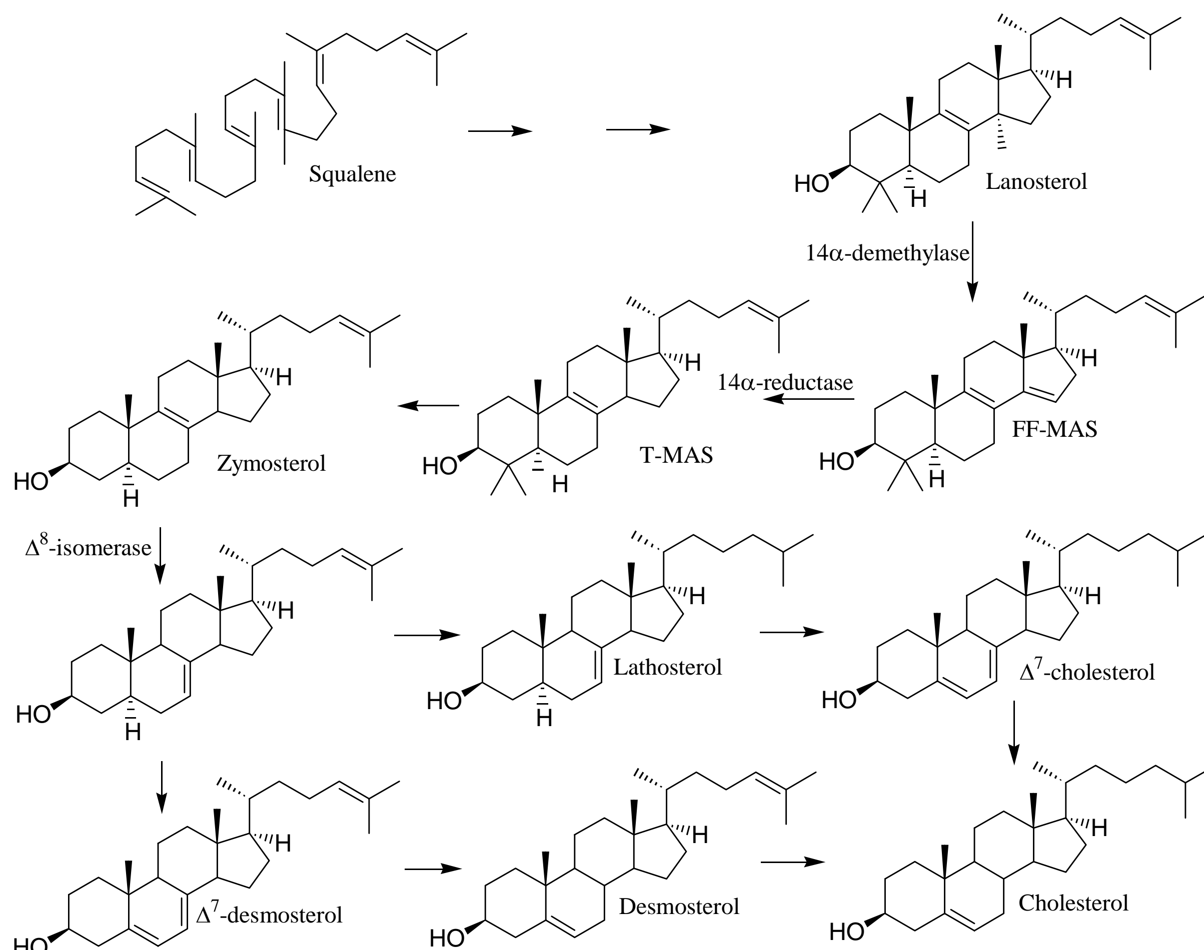


Figure 4. Schematic representation of the biosynthetic pathway of the cholesterol.

Activity of the MAS compounds

As stated above, primary oocytes contain a clearly visible germinal vesicle that breaks down when meiosis is resumed (GVBD) after the first arrest. This phenomenon is used to determine whether a compound can override the inhibitory action of hypoxanthine in *in vitro* tests. All FF-MAS related known agonists and antagonists are structurally very similar to FF-MAS (N.V. Organon, personal communication). They contain an androstane skeleton and a 3β-OH group. The 3β-OH group is mandatory, indicating that the ring A moiety plays a major role in binding MAS ligands to their receptor. All compounds contain a lipophilic aliphatic chain attached to C17. The double bond between atoms C24 and C25 enhances activity compared to compounds with a single bond at this position, but it is not mandatory (Byskov *et al.*, 1995; Strömstedt *et al.*, 1998; Wenckens *et al.*, 1998). Similarly, two methyl groups at C4 increase

the activity of the MAS ligands, but do not have to be present. Replacement of C22 by an oxygen atom leads to loss of activity for compounds with a similar configuration of C20 of FF-MAS (natural configuration). In contrast, when C22 is replaced by an oxygen atom in compounds with inverted stereochemistry of C20 (unnatural configuration), compounds can be active.

The double bond in the side chain of the MAS compounds can be mimicked by a phenyl group, as is observed in a number of structures. The same holds for the $\Delta^{8,14}$ double bond system, which can be replaced by an aromatic C ring (Grønvald *et al.*, 1997). The presence of a conjugated pi-system at ring C therefore seems mandatory. Other double bond isomers, such as Δ^{5}, $\Delta^{5,7}$, $\Delta^{7,9(11)}$ and $\Delta^{7,14}$ structures, did not result in active compounds. However, a $\Delta^{6,8(14)}$ compound was partially active.

MAS and Vitamin D3

The MAS agonists share an interesting feature in the structure-activity relationships of the stereochemistry at C20 (see above) with that of 1α,25-dihydroxy-vitamin-D_3 (calcitriol, see Figure 5) and its analogs. It was found that calcitriol analogs with an unnatural stereochemistry at C20 induced a higher transcriptional activity than those with natural configuration (Midland *et al.*, 1993; Boullion *et al.*, 1995; Yamamoto *et al.*, 1996; Liu *et al.*, 1997; Yamada *et al.*, 1998; Väisänen *et al.*, 1999; Yamamoto *et al.*, 1999). Similarly, 22-oxa compounds derived from FF-MAS with inverted C20 stereochemistry were able to induce meiosis *in vitro*.

Calcitriol is the endogenous ligand of the vitamin D receptor (VDR). The VDR is a member of the superfamily of nuclear receptors, for some of them a crystal structure is available (*e.g.* Renaud *et al.*, 1995; Wagner *et al.*, 1995; Brzozowski *et al.*, 1997; Nolte *et al.*, 1998; Tanenbaum *et al.*, 1998). It is not unreasonable to assume that the topology of the receptor active sites of the VDR and the MAS receptors are similar considering the similarity in side chain structure activity relationship (SAR). As a consequence, the C20 epimeric compounds of MAS may exhibit a binding mode similar to calcitriol and its analogs upon binding to their respective receptors. Analysis of the binding of calcitriol to its receptor may therefore increase insight into the structure-activity relationships of the C20 stereochemistry of the MAS compounds.

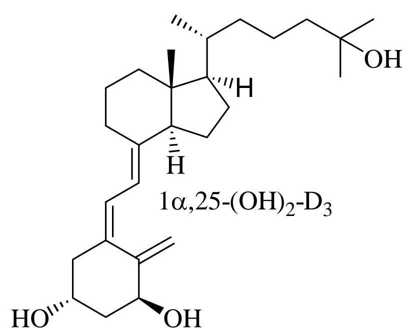


Figure 5. Chemical structure of 1α,25-dihydroxy-vitamin D_3 (calcitriol).

Methodologies

Computational modelling

Computational methods are in general inexpensive and valuable techniques for exploration of conformational behavior of molecules. Basically, two different methods can be distinguished, which are based on quantum mechanics and classical mechanics respectively. The first approach, called *ab initio* methods, use the quantum-mechanical Schrödinger equation and can in principle give an exact solution to a specific problem. Semi-empirical methods are derived from the *ab initio* method, but use parameterization to save computer time at the expense of accuracy. The second method is called molecular mechanics. It describes a molecule using classical mechanistic equations. The semi-empirical and molecular mechanics methods are discussed in more detail below.

Quantum chemical *ab initio* methods are time consuming and become increasingly expensive with increasing size of the modeled system. Most of the computer time is needed to calculate energy terms associated with electron-electron repulsion terms that occur in the Schrödinger equations for systems with more than one electron. Semi-empirical methods aim to reduce the calculation time and memory space by neglecting or approximating some of the two-electron overlap integrals that need to be calculated in the *ab initio* method. Specifically, only integrals associated with the valence electrons on the atoms in the system under consideration are evaluated. The rationale for this approach lies in the fact that most molecular properties of interest are related to the behavior of these electrons. The core electrons are considered to be part of the nuclear core of an atom. The energy contributions related to interactions between electrons in the core and the energies related to interactions between core and valence electrons are parameterized using experimental data, hence the name of the method. Several methods have been developed which differ in the degree of neglect of the two-electron integrals. The most widely used methods are MNDO, AM1 and PM3. One of the applications of the semi-empirical methods is the calculation of partial atomic charges, based on the distributions of the valence electrons on the atoms.

Molecular mechanics does not explicitly include electrons. Rather, it treats atoms as soft spheres connected to each other by a classical spring. The optimum bond distance and the magnitude of the force constant of the spring are parameters derived from experimental data. Similarly, bond and torsion angles are described by energy terms that use experiment-derived parameters, as are other non-bonded intra- and intermolecular interactions. The specific terms and parameterization used are called the force field and they determine the accuracy of the method. The energies calculated by molecular mechanics are only indirectly related to physical values, *i.e.* energy differences of different conformations indicate their relative presence in the system used in the modeling process. The fraction of a conformation present in the modeled

environment is related to the energy through Boltzmann statistics. Often, one is interested in the ensemble of conformations that a molecule can adopt. Sampling of the possible conformations can be achieved via several techniques. Generally, such an ensemble is generated by simulation of molecular movement with time, using a technique called molecular dynamics. The atomic motions (speed, acceleration) and the driving forces of the system are described by classical Newtonian equations. Another technique is stochastic sampling, in which the atom positions are modified randomly.

X-ray diffraction

X-ray crystallography is a technique to obtain experimental information on structural and conformational aspects of molecules. A successful X-ray analysis gives a structural model of the molecule(s) in the crystal in the form of the atomic positions. The method requires crystals of sufficient quality and size. A crystal can be regarded as a repetition of a building block, called the unit cell, in three dimensions. The quality of the crystal and the data obtained from it depend on the extent to which translation symmetry of this unit cell throughout the crystal is fulfilled. Obtaining good quality crystals is therefore an important part of a crystal structure analysis and has received much attention. The best crystals are usually obtained by slow growth from a slightly over-saturated solution. When the crystal is irradiated with electromagnetic waves, interaction between this incident beam and the electron cloud in the crystal cause it to scatter. As a result of the periodicity in the crystal, interference results in extinction of the scattered beam in all but a few discrete directions. These directions are indicated by indices h, k and l and depend on the geometry of the unit cell, which is defined by three translation vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. The intensities of the diffracted X-ray beams are related to the electron density distribution in the unit cell.

Fourier transformation can in principle be used to reconstruct the electron density of the unit cell from the diffraction intensities and the phase differences of the diffracted beams with the incident beam, called the structure factors. However, information on the phases is lost during the diffraction experiment. This is known as the ‘phase problem’. Two solutions to this problem have been developed in chemical crystallography. One approach, referred to as ‘direct methods’, uses strong reflections and certain combinations thereof to make an educated initial guess of the phases. In the second approach, called the Patterson method, a Fourier transform is calculated using only the observed intensities. The resulting Patterson map represents the self-convolution function of the electron density in which the position vectors of points with maximum intensity correspond to interatomic vectors in the crystal structure. The intensity of a Patterson peak is related to the atomic number of the contributing atoms. For small molecules, reconstruction of the electron density distribution is nowadays straightforward in most cases. The electron density obtained is an average of time and over the periodicity of the crystal. Since atoms move and the

periodicity of a crystal is not perfect, the electron density is somewhat smeared out over a limited area. This effect is taken into account by atomic displacement factors. The initial model is a rough estimation of the electron density and needs to be refined to increase the model sophistication. In the refinement stage, an anisotropic description of the atomic movement is introduced and disorder, if present, is handled. It is the most time consuming step of a crystal structure determination since it requires the personal attention of the crystallographer.

A large number of crystal structures have been solved so far and have been collected in rapidly extending structural databases. These include the Cambridge Structural Database (CSD, ~233.000 structures in April 2001) for small organic and organometallic molecules, the Brookhaven Protein Databank (PDB, ~15.000 structures in May 2001) for proteins, nucleic acids, DNA and other biologically interesting macromolecules. These databases contain valuable information on the conformation and intermolecular interactions between molecules. Therefore, these databases are valuable tools in the design and modeling of new materials and molecular structures.

Nuclear Magnetic Resonance spectroscopy

Where crystallographic structure determinations yield molecular models in the solid phase, high-resolution solution Nuclear Magnetic Resonance (NMR) experiments provide structural and conformational details of molecules in solution. The spin states of atomic nuclei are degenerate when an external magnetic field is absent. In the presence of an external magnetic field, the degeneracy between spin states is removed for nuclei with a quantum spin number of a ½ or higher. The energy differences associated with these spin states lie in the range of radio waves (RF). When an RF pulse is applied to the sample in the external magnetic field, resonance of the nuclear spins will occur at discrete frequencies. These resonance frequencies depend on the chemical environment around the nucleus, *i.e.* the electron density distribution, because the external magnetic field is shielded by the presence of the electronic cloud. During acquisition of the data, all these frequencies are measured at once, and the resulting collected data is referred as a FID (Free Induction Decay). After Fourier transform of this FID, the NMR spectrum is obtained, in which the intensity of the signal is given as function of the resonance frequency of the nucleus. The latter is often given in parts per million (ppm), also referred to as chemical shift, which is independent of the NMR spectrometer used. Magnetization on spins can be transferred through bonds, which gives rise to a phenomenon called *J*-coupling. The result of coupling in the NMR spectrum is splitting of the signal of a spin system.

NMR spectroscopy is widely used as an analytical tool for identification of small molecular compounds. For this purpose, information on the chemical shifts and coupling constants usually suffice, which can be obtained through one-dimensional NMR experiments as described above. Derivation of more

detailed structural information on molecules requires the use of more sophisticated NMR techniques. In general, multi-dimensional NMR experiments are applied for these purposes. The discussion here will focus on 2D experiments, with two frequency axes. A two-dimensional NMR experiment involves four stages. During the first stage, the preparation stage, magnetization is transferred to a certain spin. During the following evolution stage, chemical shift evolution occurs over the period of the first time variable (t_1). By incrementing t_1 in a series of measurements, modulation of the chemical shifts occurs, which enables the detection of the various nuclei in the second dimension of the 2D spectrum. The third stage is used to manipulate the spin system in some way, the details depend on the type of experiment. In the experiments described below, magnetization is transferred from one spin to another, which is called mixing. During the final detection stage, chemical shift evolution occurs over the duration of the second time variable (t_2) and magnetization on the labeled spins is detected. A two-dimensional matrix of resonance data is acquired as a result, which gives the 2D spectrum after Fourier transformation in both dimensions. Magnetization transfer during the evolution time can be achieved in various ways, leading to different types of experiments. Two of those will be discussed here, called correlated spectroscopy (COSY) and nuclear overhauser enhancement spectroscopy (NOESY).

The key concept of the COSY methods is transfer of magnetization from one spin to another via ^3J coupling, which occurs through bonds. The transfer results in cross peaks that appear above and below the diagonal line of the 2D spectrum at the intersections of the frequencies of the two spin systems involved. From the cross peaks, atom connectivity's can be deduced and used to identify the molecular structure. In NOESY techniques and its variants, the principle of magnetization transfer is comparable, but the mechanism differs: nuclei can exchange magnetization via cross-relaxation, which occurs through space. The cross peaks are associated with nuclei at a short distance r, which is usually smaller than 5 Å, and their intensities are proportional to r^{-6}. From the intensities of the cross peaks, the inter-nuclear distances can be approximated. These distances are averages over time and over the ensemble of conformations present in the solution. From these distances the conformation or conformations of a molecule can be derived.

Homology modeling

For most proteins, experimental information on the structure is not available. Some of these proteins are involved in mechanisms that are associated with diseases. These proteins constitute interesting targets for the development of new drugs and structural information is often of great assistance in the development of new pharmaceutical agents. To remedy the fact that the structure of such a protein is not known, a structural model can be obtained if structural information is available on proteins that are evolutionary related. Related

proteins often have a high percentage of identical or similar residues in their sequences. Since the fold of a protein is largely determined by the amino acid sequence, one can assume that proteins with a high sequence homology have similar folds. A sequence of the protein with known structure can then be used as a template onto which the residues of the protein with unknown 3D structure are matched. This process, called 'alignment', becomes increasingly prone to errors with decreasing homology. As a consequence, the model decreases in quality when the homology decreases. After optimal alignment, at least forty percent of the residues should be identical to those of the template structure to ensure a reliable model.

The homology model is built by first constructing the backbone of the model protein using the coordinates of the template protein, excluding the loops between α-helices and β-sheets. Thereafter, the side chains of identical residues in the sequence alignment are included with similar conformation as in the template. The side chains that do not match those of the template are constructed using statistical information on side chain conformations present in the PDB database (see above). The loops present the largest challenge in the construction of a homology model because deletions, insertions and mutations occur most often there. Several methods have been developed to model these loops, which will not be discussed here. Often, several protein models are built and one is selected after the model is complete. Finally, the selected protein model is minimized using a mild procedure, to relieve strain caused by atomic clashes. A homology model can yield valuable information on potential binding modes of molecules and on structural and conformational requirements of ligands.

SuperStar

The design of new drugs is helped greatly if the tertiary structure of the target protein has been solved. The design of new, potent ligands using the protein structure requires a good understanding of molecular interactions and an adequate model to describe them. Several approaches have been developed with this goal. The most widely used methods are based on molecular mechanics. These methods calculate energies of interaction for a given system of molecules and evaluate the binding potency of a ligand based on these energies. Other methods have been developed which rely more directly on the experimental data on molecular interactions present in structural databases (Danziger & Dean, 1989a,b; Böhm, 1992a,b; Clark *et al.*, 1995; Laskowski *et al.*, 1996). One such a method is SuperStar (Verdonk *et al.*, 1999, 2001).

SuperStar locates favorable interaction sites of a chemical functional group near a molecule. The method uses intermolecular interaction information stored in a database called IsoStar (Bruno *et al.*, 1997, 1999). The IsoStar database consists of geometrical distributions of one chemical group, called the contact group, around another, called the central group. The distributions are derived from the crystal structures of the CSD and protein-ligand complexes from the

PDB. For a SuperStar run, a set of distributions with identical contact groups is used. The central groups of the distributions are superposed on matching molecular fragments in the molecule, thereby transforming the distribution of the contact group accordingly. After matching of all structural fragments, the studied molecule is surrounded by the separate contact groups distributions, some of which overlap. In the final step, SuperStar combines the separate distributions to give a composite propensity distribution for a contact group. From these composite distributions, favorable interaction sites for a contact group in the vicinity of a molecule (*e.g.* in a protein active site) can be identified. These hot-spots can then be used to find the binding mode of a known ligand and to design new ligands.

Outline of this thesis

The main purpose of the work described in this thesis is the improvement of the understanding of MAS structure and conformation, in particular of FF-MAS and its analogs. In addition, the relative stabilities of isomers of MAS analogs are studied to gain insight into the synthesis of these compounds. Heats of formation, based on molecular mechanics methods, are calculated for this purpose. This work is reported in chapter 2. Structural and conformational aspects of the side chains in the MAS structures are studied using crystallographic structure determinations and NMR spectroscopy. In chapter 3, the crystal structures of several active and inactive MAS analogs are presented and the conformations and their implication on structure-activity relationships are discussed. In addition, the conformations of the side chains of three MAS analogs in solution are discussed in this chapter. In addition to the conformation, the molecular electrostatic potential of several double bond isomers is compared using semi-empirical calculations. The results of these calculations are reported in chapter 4.

In an attempt to exploit the analogy in the SAR of side chains of FF-MAS and calcitriol, a homology model of the VDR was built, to which calcitriol and its C20 epimer were docked. SuperStar is particularly suited for this purpose. The set of geometrical distributions used in SuperStar were initially derived from CSD structures. It can be argued that the nature of non-bonded interactions in crystal structures of small molecules may differ from the way non-bonded interactions occur in a protein environment, as has been suggested by Verdonk *et al.* (1999). Therefore, the method may benefit from the use of interactions between proteins and ligands as source data taken from the Brookhaven Protein Database (PDB) and present in the IsoStar database. The implementation and validation of SuperStar based on PDB data is described in chapter 5. The homology model of the VDR and the docking results by SuperStar are described in chapter 6.