Eur J Med Chem (1997) 32, 869–879

Structure–activity relationships of steroids with mineralocorticoid activity

G Grassy[1]*, J Fagart[2], B Calas[1], M Adenot[1], ME Rafestin-Oblin[2], G Auzou[3]

[1]*Centre de Biochimie Structurale, CNRS UMR 9955, INSERM U414, Faculté de Pharmacie, Université de Montpellier 1, 15, avenue Charles Flahault, 34060 Montpellier Cedex, France*
[2]*INSERM U246, Faculté de Médecine X, Bichat, BP 416, 16, rue H Huchard, 75870 Paris Cedex 18, France*
[3]*INSERM U300, Faculté de Pharmacie, 15, avenue Charles Flahault, 34060 Montpellier, France*

(Received 3 March 1997; accepted 29 May 1997)

Summary — Fourteen steroid homologues, belonging to the series of 18-substituted progesterone and 17-hydroxymethylketone derivatives were modeled by both molecular and quantum mechanics. We have studied the dependency of the affinity of these compounds for the hMR (human mineralocorticoid receptor) by means of various parameters describing the structure and its molecular properties. Using variable mapping coupled to a discriminant analysis, this work demonstrates the non-linear relationships between affinity and some structural features. We have constructed a model allowing us to predict the affinity and the activity of new compounds. The principal electronic and structural characteristics leading to a selective affinity and activity were revealed.

structure–activity relationship / variable mapping / steroids / antimineralocorticoids

Introduction

The first step in the action of aldosterone is its binding to an intracellular receptor, the mineralocorticoid receptor (MR). This receptor belongs to the nuclear receptor superfamily that includes the steroid, thyroid and retinoic acid receptors [1, 2]. These receptors share a common structural organization. They contain highly conserved DNA binding domains, a less conserved C-terminal hormone binding domain and a N-terminal region, which varies most, both in sequence and in length [3].

MR acts as a ligand-inducible transcriptional regulator controlling the activity of specific gene networks [1]. Ligand binding and the ligand-induced transactivation of MR are separated by a cascade of events including a conformational change, dissociation of the 90 kDa heat shock protein (hsp90), dimerization and binding to the hormone response element [4–11]. The change in receptor conformation is likely a key step in the receptor activation since agonists and antagonists differentially modify the MR conformation [10, 11].

The most widely used and certainly the best-documented aldosterone antagonist agents are spirolactones, which are steroidal drugs that contain at the C-17 position a γ-lactone (eg, spironolactone) or a γ-hydroxy acid moiety (eg, potassium canrenoate). It has been well established that the lactone ring is of crucial importance for MR binding and is responsible for the in vivo antagonist properties [12–14]. Extensive molecular steroid modifications, especially at the C 7, 11, 15 and 16 positions, have been shown to modulate the antagonist potency of spirolactones and their mineralocorticoid specificity, while leaving their agonist/antagonist profile virtually unchanged [14, 15] Progesterone (P) binds to MR with a high affinity and acts as an antagonist derivative [16–19]. 18-substituted progesterone derivatives have been recently synthesized [20–23]. They are characterized by an unsaturated side chain, an enone group or a diazomethyl ketone function. Contrasting to what was observed for spirolactones the agonist/antagonist properties of the 18-substituted progesterone derivatives are dependent upon the nature of the substituent [23, 24]. To point out the molecular characteristics necessary to obtain derivatives with high affinity for MR and high mineralocorticoid activity, we have analyzed the structure–activity relationships by focusing on mineralocorticoids and glucocorticoids that are

*Correspondence and reprints

characterized by a hydroxymethylketone group at the C-17 position and on 18-substituted progesterone derivatives. The aim of this study is the use of rational models that relate the physicochemical and structural properties of steroidal derivatives to their activities and affinities for the MR.

Material and methods

Molecular data

In order to study the influence of various steroid moieties on hMR binding, we recently synthesized a set of 18-substituted corticoids. Besides these products, we have also selected as a reference some typical compounds belonging to the mineralocorticoid and glucocorticoid series. All molecules of this work are summarized in figure 1.

Synthesis

18-vinylprogesterone (18VP) and 18-ethynylprogesterone (18-EP) were synthesized as previously described [20, 21] from 3β-hydroxy-pregn-5-ene-18,20-dione with suitable protections at the C-3 and C-20 positions. After reduction to the corresponding aldehyde, Wittig reaction using ϕ_3PCH_2 or ϕ_3PCHCl, followed by base-promoted HCl elimination led to the corresponding unsaturated compounds that were converted by classical methods to 18VP and 18EP respectively. 18-oxo-18-vinylprogesterone (18OVP) was generated from 18,20-epoxy-pregn-4-ene-3,18-dione [22]. 18-(diazomethyl)-20-hydroxypregn-4-ene-3,18-dione (18DOP1) and 18-(diazomethyl)pregn-4-ene-3,18,20-trione (18DOP2) were synthesized from 3β,20β-dihydroxy-18-methylpregn-5-ene-18-one [23]. The methyl ketone was converted to the formyl derivative and the diazo group was introduced by reaction with tosylazide. The oxydation at C-3 led to 18DOP1 that was converted to 18DOP2 by a Swern oxydation. 21-diazoprogesterone (21DP) was obtained by reaction of diazomethane with 3-oxoandrost-4-ene-17β-carboxylic acid chloride generated from the corresponding acid. Non-radioactive aldosterone (A), progesterone (P), corticosterone (B), cortisol (F), dexamethasone (Dex), deoxycorticosterone (DOC) and triamcinolone acetonide (TA) were purchased from Sigma Chimical Co (St Louis, MO). [1,2-^{3}H]A (40–60 Ci/mmol), [1,2-^{3}H]Dex (40–60 Ci/mmol) and [1,2,6,7-^{3}H]P (80–110 Ci/mmol) were purchased from the Radiochemical Center, Amersham (Aylesbury, Buckinghamshire, UK). [18',18''-^{3}H]18-vinylprogesterone ([^{3}H]18VP, 40–50 Ci/mmol) was synthesized as described by Delorme et al (manuscript in preparation).

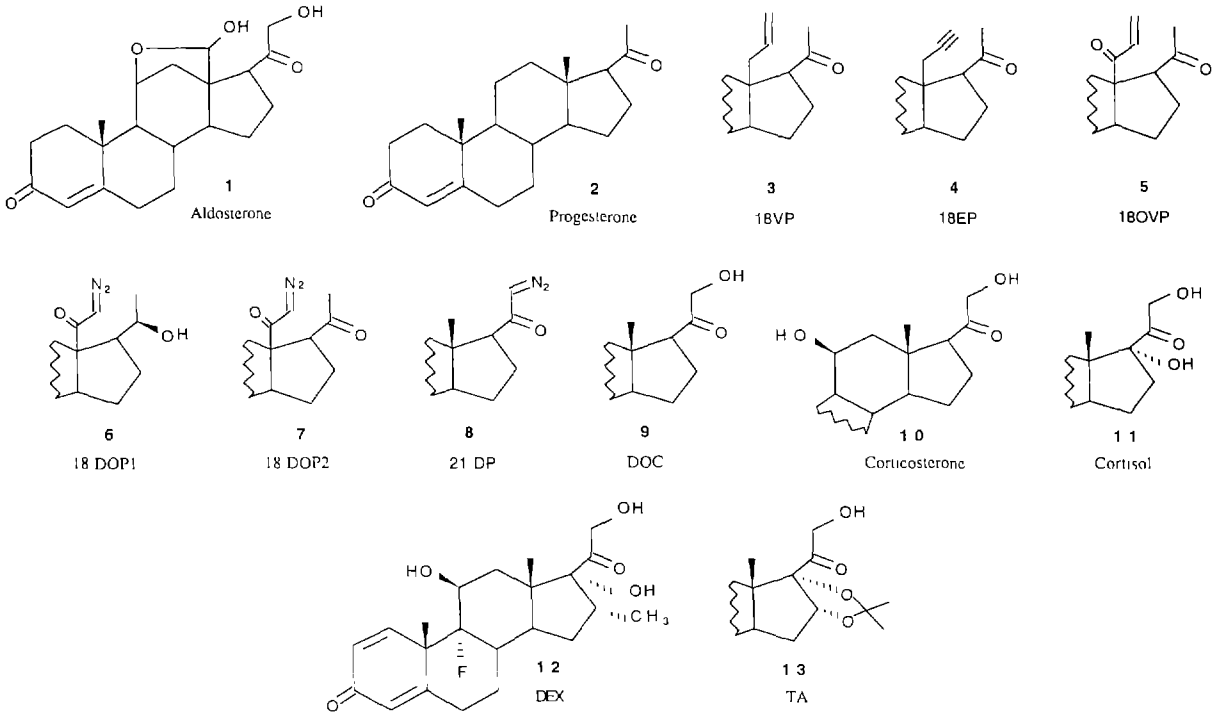
Pharmacology

Steroid binding characteristics

The affinity of the steroids was determined by Scatchard plot analysis when the tritiated compounds were available. Cytosol from Sf9 cells infected with the recombinant AcNPV-hMR [24] was incubated for 4 h at 4 °C with increasing concentrations (0.5–500 nM) of [^{3}H]steroid. Bound (B) and unbound (F) steroids were separated by the charcoal-dextran technique previously described [4]. The evolution of B as a function of F was analyzed and the dissociation constant at equilibrium (Kd) was calculated by a previously established computer method [25]. When only unlabeled steroids were available the affinity constants were measured by competition experiments using [^{3}H]aldosterone as reference. Cytosol from Sf9 cells infected with the recombinant AcNPV-hMR [26] was incubated for 4 h at 4 °C with 3 nM [^{3}H]aldosterone both in the absence and presence of increasing concentrations of unlabeled competitors (0.3–600 nM). Bound and unbound steroids were separated by the charcoal-dextran technique previously described [4]. The apparent dissociation constants at equilibrium (Kd_{app}) were calculated according to the formula $Kd_{app} = (Kd)_A \times [X]_{50}/[A]_{50}$ where $[X]_{50}$ and $[A]_{50}$ are the concentrations of the competitor and of aldosterone that induce a 50% displacement of [^{3}H]A binding and where $(Kd)_A$ is the mean value of the dissociation constant of aldosterone as determined in a previous study [25].

Mineralocorticoid activity

The mineralocorticoid activity of the derivatives was analyzed in CV1 cells, grown in minimum essential medium – MEM, by using the previously described *cis–trans* cotransfection assay [23, 24]. Briefly, CV1 cells were transfected using the calcium phosphate precipitation method [27] with 5 μg of pRShMR as expression vector for the hMR, 5 μg of pFC31Luc, which contains the MMTV promoter driving the luciferase gene, 5 μg of pCH110, which contains the gene coding for the β-galactosidase enzyme, and 5 μg of pSP72 as plasmid carrier. Sixteen hours after transfection the cells were rinsed twice with PBS and then incubated with MEM containing the tested steroid. After 24 h incubation the cells were harvested and the cell extracts assayed for luciferase [28] and β-galactosidase activity [29]. To standardize for transfection efficiency, the relatively light units, obtained in the luciferase assay, were divided by the optical density obtained in the β-galactosidase assay. All the biological results are gathered in table I.



18VP 18-vinylprogesterone, 18EP 18-ethynylprogesterone, 18OVP 18-oxo-18-vinylprogesterone, 18DOP1 18-(diazomethyl)-20-hydroxypregn-4-ene-3,18-dione, 18DOP2 18-(diazomethyl)pregn-4-ene-3,18,20-trione, 21DP 21 diazoprogesterone, DOC deoxycorticosterone, DEX dexamethasone, TA triamcinolone acetonide

Fig 1. Formulas of test steroids related to table I.

Molecular modeling

All initial modeling was done on a FDDI connected cluster of Hewlett Packard workstations APOLLO 735 and on the RHEA computer of the CNUSC Montpellier (France). Molecular structures were generated by MAD [30] and their lowest energy conformations were jointly determined by molecular mechanics (MM_2 force field coupled to steepest descend and Newton–Raphson minimisers) and by semi-empirical quantum mechanical methods (MOPAC 6.0 using the AM_1 Hamiltonian and the options PRECISE [31]).

The molecular electrostatic potentials (MEPs) and molecular similarity indices were determined by the programme ASP [32], based on the flexible conformations compared to the aldosterone reference [33].

Steroids are nearly rigid structures, so we have determined the lowest-energy conformations using both a semi-empirical quantum mechanical method and molecular mechanics (in vacuo). Validation of the results was done by evaluation of the RMS between

Table I. Mineralocorticoid affinities and activities of test-substances.

Compounds	*Affinities (Kd nM)*	*Activities (IC_{50}[a], ED_{50}[b] M)*
1[c]	1	10^{-11} (antagonist)[b]
2	0.39/0.47	10^{-8} (antagonist)[a]
3	0.86/1.79	10^{-8} (antagonist)[a]
4	4.95	10^{-7} (antagonist)[a]
5	105	10^{-7} (agonist)[b]
6	20	10^{-6} (antagonist)[a]
7	20	10^{-6} (antagonist)[a]
8	12	10^{-7} (agonist)[b]
9	0.3	ND (agonist)
10	0.75	10^{-7} (agonist)[b]
11	0.7/0.5	$5 \cdot 10^{-9}$ (agonist)[b]
12	1	10^{-6} (agonist)[b]
13	0.6	ND (agonist/antagonist, 50:50)

[c]The formulas of the compounds are presented in figure 1.

the conformation generated by those means and the corresponding X-ray structure available for certain steroids.

Structure–activity relationships

Structure–activity relationships were carried out with the TSAR 2.31 [34] software implemented on a SGI Indigo Elan R4000 graphics workstation.

Molecular properties descriptors

Using the molecular conformations we have generated a set of physico-chemical parameters that are representative of structural, electronic and lipophilic properties of the compounds (table II).

Topological descriptors
This kind of descriptor is used to quantitatively describe the molecular topology. We selected Balaban [35], Wiener [36], Randic [37] and Kier indices [38] and the 2D autocorrelation vector [39].

Electronic parameters
In order to characterize the relevant electronic properties for the electrostatic interaction with the receptor, we have selected the electronic densities carried by all the oxygen atoms, the global molecular dipole (MOPAC), the first ionisation potential, the Heat of formation (kcal/mol) and also the HOMO and LUMO energies.

Structural parameters
The overall molecular form is described by a set of 3D auto-correlation vectors [39] calculated with a step of 1 Å and weighted by different atomic properties. We computed both molecular volumes, and accessible molecular surfaces and the molecular inertia momentum components I_x, I_y, I_z.

Based on the interdistance matrices, the distances O_1–O_2, and maximal distance between O atoms ($d_{O_{max}}$) were also chosen as representative for the structural variability of the compounds (fig 2).

Lipophilicity parameters
The lipophilicity was evaluated by the determination of Log *P* using the atomic incremental method of Ghose and Crippen [40]. The distribution of the molecular lipophilicity all along the structure is represented by 3D autocorrelation vectors weighted by atomic lipophilic increments. We also computed the molar refractivity (MR).

Molecular similarity indexes
Structural details of the hormone binding site of hMR are not yet available, and, for our team, several attempts to crystallise the protein corresponding to the ligand binding domain (LBD) of hMR were unsuccessful. Consequently, direct docking studies based on 3D models of the receptor are hitherto irrelevant.

Despite this fact, the binding mode of a drug to its receptor is dependent on the intermolecular forces between them. The drug's molecular shape, the spatial distribution of lipophilic and electrostatic properties will determine how it interacts with the receptor site. Molecular similarity calculations using the natural ligand (aldosterone) as a reference may be expected to correlate with biological activities. Using the ASP software we have computed three molecular similarity indices using a GRID based Carbo [32] calculation of Gaussian curves reproducing:
- the molecular electrostatic potential (Carbo electrostatic),
- the molecular lipophilic potential (Carbo lipophilic),
- the electron density functions [32] for different atom types (derived from STO3G atomic orbitals) and estimating the shape similarity (Carbo shape).

For these three properties, the optimisation of the superimposition of molecules was done both by translation, rotation and flexible fitting, using Aldosterone as a rigid reference. These indices quantify the similarities of the spatial potentials (shape, lipophilicity and electrostatic) induced by aldosterone and the other molecular structures of this study.

Statistical methods

For the structure–activity relationships, the statistical methodology used was:
- linear multivariate methods: multiple stepwise regression (MSR), principal component analysis (PCA), stepwise discriminant analysis (SDA);
- nonlinear methods: variable mapping and cluster significance analysis (CSA).

Principal component analysis
This is a multidimensional statistical method [41] for data analysis well suited for representation of molecules in the hyperspace of their properties (molecular descriptors). The PCA can be used to reduce a large number of descriptors to a smaller number of synthetic orthogonal variables resulting from a linear combination of these original descriptors. This method conserves the largest part of the total initial information. The original variables were normalised, the diagonalization of the covariance matrix was done using the classical Jacobi transform routine.

Stepwise discriminant analysis
This method [41] attempts to produce a qualitative classification (eg, inactive and active molecules) using a linear combination of numerical descriptors. This

method is derived from PCA as well as from multiple regression. Discriminant analysis determines a classification rule which can be used to predict the class membership of unknown compounds.

Cluster significance analysis

Cluster analysis [42] is a technique used to group a set of points into groups that consist of similar members, based on their distances in a choosen parameter hyper space. If the starting clusters are fixed by experiments (typically active or inactive), it is possible to test the validity of a parameter hyperspace by using an algorithm that compares the validity of the proposed a priori classification and the resulting distribution of distances in that space defined by their descriptors. In our case, the significance of the classification is calculated by testing all the possible combinations of individuals.

The validity of the classification is estimated from the numerical value of a probability, the analysis being more relevant as the probability approaches 1.0. For the probabilities close to 0.5 the proposed classification is not more valid than those obtained by chance in parameter space. This method is particularly useful in the following cases [43]:

(1) when there is no linear relationship between activity and explanatory parameters;

(2) when only one narrow variation interval of properties leads to an interesting biological activity.

This method is complementary to the analysis of the properties by variable mapping.

Variable mapping

This qualitative technique [44] consists of an evaluation of the distribution (global or percent wise) of the active and inactive molecules as a function of the distribution of parameter values. The superposition of the ensemble of graphs (activity property) is supposed to indicate, for certain parameters, the limiting values (inferior or superior) necessary for activity. This graphical method, which is simple and rapid, gives a diagnosis of the qualitative non-linear dependencies between the activity and a molecular property. For the properties that are involved in receptor ligand interactions, it has been clearly established that the existence of strict contingencies that determine the adaptibility to the receptor imply an embedding of certain structural and physico-chemical poperties. This method determines simple rules which can be used to predict the activity of unknown products. A graphical representation showing the number of successes relative to the number of violations of the rules allows comparison of the distributions with the activities for the ensemble of molecules under study.

The validity of the representation is estimated from the efficiency of the prevision P defined as 100 times the ratio between the number of molecules possessing a certain activity whose properties correspond to the proposed distribution and the total number of molecules obeying the same distribution criteria. A subroutine allowing a systematic diagnosis of the variables with 'embedding' of the structural, topological or physico-chemical properties, generating the corresponding graphs and ratios, was implemented in our software TSAR.

Results

Analysis of the affinity for the human mineralocorticoid receptor

A dual classification was taken into account, dividing the molecules into two groups:

- the group of molecules that have high affinity ($\leq$ 1 nM): Progesterone, Aldosterone, 18VP, Corticosterone, Cortisol, Dexamethasone, Deoxycorticosterone (DOC) and Triamcinolone (TA) (high affinity group);
- the group of molecules that have low affinity ($>$ 1 nM): 18EP, 18OVP, 18DCP1, 18DCP2, 21DP (low affinity group).

After a selection based on the correlation matrices (in order to avoid redundancies) we have kept a limited number of explicative parameters (table III).

Classical statistical analysis

Stepwise regression. The formalism of linear free energy relationships involved in QSAR analysis suggest that ΔG, the free energy of binding, is proportional to the binding affinity $-\log K$, so we expressed the biological activities of steroids as $\log_{10}(1/C)$ expressed in nM.

No significant linear dependency, expressed as a regression equation, exists between the affinity towards hMR expressed as $-\log_{10}C$ and the set of molecular descriptors.

The best regression equation is obtained with the Kappa 3 index of Kier:

$-\log_{10}C = -2566$ Kappa 3 + 15.30,

$r^2 = 0.520$, $F = 11.91$, $s = 0.584$, but the cross validated $r^2 = 0.216$.

As is frequently observed, the lack of significance of this regression analysis may be due to a non linear dependency between the biological response and the molecular descriptors.

Principal component analysis. A principal component analysis performed on 3 components of the 2D autocorrelation vector (Bin 2, Bin 10 and Bin 13) in order to avoid the redundancies, shows the localisa-

Table II. Structural descriptors.

Ref number	*Name*	*HOMO energy level (eV)*	*LUMO energy level (eV)*	*Heat of formation (kcal/mol)*	*Ionisation potential (eV)*	*Distance O_1–O_2*	*$d_{O_{max}}$*
1	Aldosterone	–10.203	–0.159	–222.833	10.203	11.677	12.463
2	Progesterone	–10.069	–0.040	–106.887	10.069	11.657	11.657
3	18 VP	–10.071	–0.044	– 85.659	10.071	11.701	11.701
4	18 EP	–10.087	–0.056	– 47.886	10.087	11.738	11.738
5	18 OVP	–10.037	–0.014	–106.749	10.037	11.779	11.779
6	18 DOP1	– 9.161	–0.014	– 82.601	9.161	11.344	11.344
7	18 DCP2	– 9.376	–0.137	– 64.853	9.376	11.789	11.789
8	21 DP	– 9.400	–0.255	– 35.968	9.400	11.642	11.642
9	DOC	–10.105	–0.070	–153.084	10.105	11.676	12.311
10	Corticosterone	–10.145	–0.084	–193.478	10.144	11.669	12.264
11	Cortisol	–10.189	–0.123	–233.041	10.189	11.641	12.337
12	Dexamethasone	–10.177	–0.467	–243.039	10.177	11.625	12.323
13	Triamcinolone	–10.083	–0.363	–263.029	10.083	11.713	12.441

Ref number	*Charge O_1*	*Charge O_2*	*Total dipole (Debye) VAMP*	*Carbo index Gaussian combined Aldosterone*	*Carbo index Gaussian electrostatics Aldosterone*	*Carbo index Gaussian shape Aldosterone*	*Carbo index Gaussian lipophilicity Aldosterone*
1	–0.497	–0.471	1.602	1	1	1	1
2	–0.501	–0.458	2.903	0.729	0.481	0.749	0.959
3	–0.499	–0.463	2.704	0.859	0.701	0.917	0.959
4	–0.5	–0.465	2.524	0.894	0.755	0.961	0.966
5	–0.503	–0.467	4.638	0.865	0.667	0.966	0.964
6	–0.501	–0.498	4.569	0.696	0 341	0.794	0.954
7	–0.502	–0.456	5.185	0.751	0.506	0.777	0.97
8	–0.5	–0.505	2.855	0.667	0.222	0.814	0.966
9	–0.499	–0.472	2.088	0.924	0.806	0.986	0.978
10	–0.498	–0.475	2.901	0.928	0.82	0.973	0.991
11	–0.496	–0.489	3.016	0.892	0.715	0.971	0.991
12	–0.52	–0.496	4.802	0.803	0.494	0.926	0.99
13	–0.521	–0.491	4.614	0.622	0.149	0.789	0.927

Ref number	*Log P (whole molecule)*	*Total lipole (whole molecule)*	*Molecular volume (whole molecule)*	*Surface area (whole molecule)*	*Molar refractivity (whole molecule)*	*Randic index (H excluded) (whole molecule)*	*Balaban index (H excluded) (whole molecule)*
1	1.583	4.00E–06	273.150	308.57	95.39	12.41	1.29
2	3.689	2.00E–06	256.828	306.73	92.80	10.86	1.35
3	4.266	2.00E–06	279.515	322.05	102.04	11.92	1.39
4	3.800	1.00E–06	275.925	316.69	100.46	11.92	1.39
5	3.283	3.00E–06	286.082	330.77	102.81	12.34	1.42
6	3.228	1.00E–06	291.368	339.73	104.76	12.84	1.43
7	3.437	3.00E–06	289.805	330.31	103.73	12.84	1.43
8	3.658	1.00E–06	272.150	315.80	98.46	11.90	1.53
9	3.133	2.00E–06	265.542	312.09	94.50	11.40	1.35
10	2.148	5.00E–06	274.557	315.58	96.09	11.81	1.39
11	1.431	3.00E–06	282.222	318.73	97.49	12.17	1.43
12	1.712	3.00E–06	300.124	325 66	102.59	12.96	1.52
13	1.921	2.00E–06	328.299	355.75	111.68	14.35	1.38

Table II. *Continued.*

Ref number	*Wiener index (H excluded) (whole molecule)*	*Randic index (H included) (whole molecule)*	*Balaban index (H included) (whole molecule)*	*Wiener index (H included) (whole molecule)*	I_1 *(whole molecule)*	I_2 *(whole molecule)*	I_3 *(whole molecule)*
1	1401	23.421	1.960	7380	153.779	746.167	819.836
2	1052	22.592	2.264	7025	120.334	628.013	685.737
3	1316	24.446	2.329	8523	163.790	681.081	743.663
4	1316	23.653	2.293	7769	157.620	683.059	750.278
5	1441	24.107	2.291	8217	173.888	714.267	758.090
6	1614	24.640	2.324	8537	211.422	736.384	797.073
7	1614	23.734	2.280	7817	204.052	768.504	819.716
8	1374	22.829	2.581	7075	165.543	723.810	804.573
9	1201	23.153	2.265	7439	136.159	721.367	795.189
10	1305	23.713	2.296	7748	153.578	728.904	806.526
11	1425	24.274	2.321	8089	165.062	769.224	836.159
12	1670	24.845	2.356	8456	196.812	806.806	880.208
13	2185	26.949	2.058	10757	259.312	938.851	1010.984

Ref number	R_1 *(whole molecule)*	R_2 *(whole molecule)*	R_3 *(whole molecule)*	*Ellipsoidal volume (whole molecule)*	*E-state sum (whole molecule)*	*Kappa 1 (whole molecule)*	*Kappa alpha 1 (whole molecule)*
1	17.536	3.614	3.289	873.143	61.000	18.056	17.085
2	16.734	3.206	2.937	660.032	46.833	16.467	15.599
3	18.329	4.408	4.037	1366.267	51.333	18.367	17.245
4	18.636	4.300	3.915	1314.358	52.833	18.367	17.074
5	17.448	4.248	4.002	1242.449	58.500	19.322	17.882
6	14.220	4.083	3.772	917.150	63.167	20.280	18.768
7	14 718	3.908	3.664	882.603	64.500	20.280	18.491
8	15.451	3.534	3.179	727.144	55.833	18.367	16.903
9	18.180	3.431	3.113	813.426	52.333	17.416	16.505
10	16.619	3.501	3.164	771.332	58.167	18.367	17.416
11	16.060	3.446	3.170	735.017	64.083	19.322	18.329
12	16.467	4.017	3.682	1020.200	74.833	21.240	19.925
13	16.236	4.484	4.164	1269.969	79.083	22.776	21.436

Ref number	*Kappa 2 (whole molecule)*	*Kappa alpha 2 (whole molecule)*	*Kappa 3 (whole molecule)*	*Kappa alpha 3 (whole molecule)*	*Flexibility (phi) (whole molecule)*	*Total dipole (whole molecule)*
1	5.998	5.491	2.234	2.006	3.608	1.594
2	5.500	5.037	2.289	2.057	3.416	2.890
3	6.558	5.926	2.667	2.354	4.087	2.727
4	6.558	5.831	2.667	2.308	3.982	2.306
5	6.805	6.003	2.704	2.317	4.128	4.482
6	7.356	6.493	2.971	2.547	4.513	3.675
7	7.356	6.338	2.971	2.473	4.341	5.633
8	6.558	5.737	2.749	2.335	3.879	2.708
9	6.021	5.520	2.481	2.232	3.797	1.944
10	6.270	5.750	2.588	2.329	4.006	2.735
11	6.250	5.728	2.486	2.236	4.038	3 047
12	6.498	5.833	2.395	2.100	4.151	4.622
13	6.780	6.129	2.606	2.308	4.238	4.489

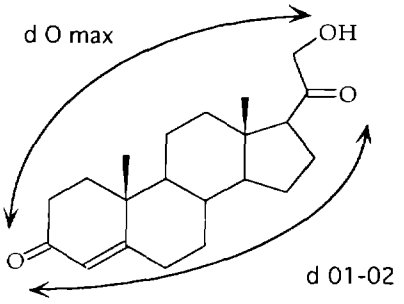


Fig 2. Definition of $d_{O_1-O_2}$ and $d_{O_{max}}$ distances.

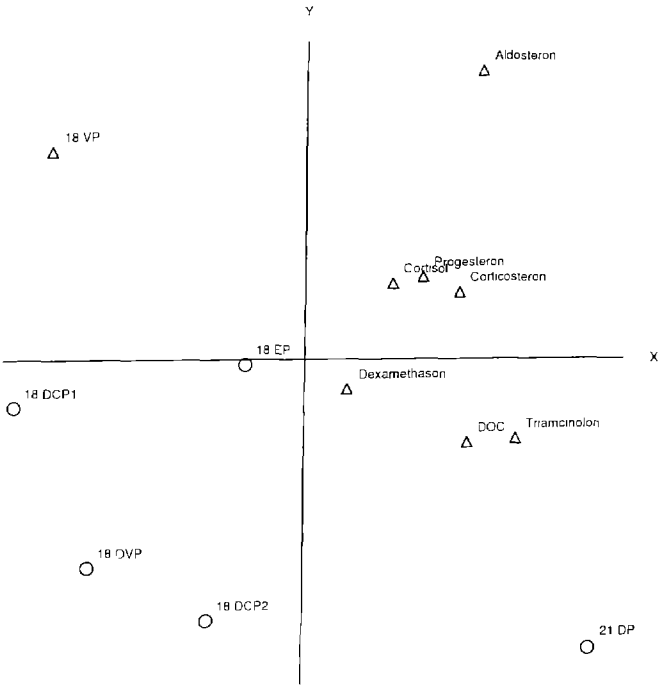


Fig 3. Principal component analysis of hMR ligands using 2D autocorrelation vectors. Total variance explained in this plane, 93.6%; triangles: high-affinity compounds; circles: low-affinity compounds.

tion of high-affinity compounds in a narrow area in the principal plane defined by the two first principal components (fig 3) and explains 93.6% of the total variance. This analysis demonstrates the dependancy between the biological affinity towards the hMR and the shape of the ligands described by the topological 2D VAC.

Variable mapping and cluster significance analysis

In order to detect an eventual embedding of certain properties, a variable mapping was done on the whole set of previous described parameters, taking into account the qualitative distributions of the affinity for the hMR receptor as defined. The results are gathered in figure 4. The structural and physico-chemical constraints, deduced from this study and necessary for high affinity, are:

Topological constraints. The topological constraints belong from chemical structures and are strongly dependent of 2D molecular graph. For some topological indices, we observed significant differences on the assigned values of high-affinity compounds vs low-affinity ones: a low Kappa 3 index from Kier (< 2.65) as well as a low Kappa $\alpha 3$ (< 2.35) and a low flexibility (Phi < 4.25).

Table III. Correlation matrix of parameters.

Var	X_1[a]	X_2[b]	X_3[c]	X_4[d]	X_5[e]	X_6[f]	X_7[g]	X_8[h]
X_1	1.0	–0.64	–0.59	0.014	0.614	–0.721	–0.586	–0.507
X_2		1.0	0.619	–0.456	–0.690	0.80	0.719	0.501
X_3			1.0	0.002	–0.891	0.599	0.512	0.061
X_4				1.0	0.303	–0.166	–0.230	–0.263
X_5					1.0	–0.622	–0.555	–0.217
X_6						1.0	0.971	0.791
X_7							1.0	0.799
X_8								1.0

[a]X_1, $-\log_{10}C$ molar affinity; [b]X_2, HOMO energy level (eV); [c]X_3, heat of formation (kcal/mol); [d]X_4, distance O_1–O_2 (Å); [e]X_5, $d_{O_{max}}$; [f]X_6, Kappa 3 (Kier); [g]X_7, Kappa α_3; [h]X_8, flexibility (Phi).

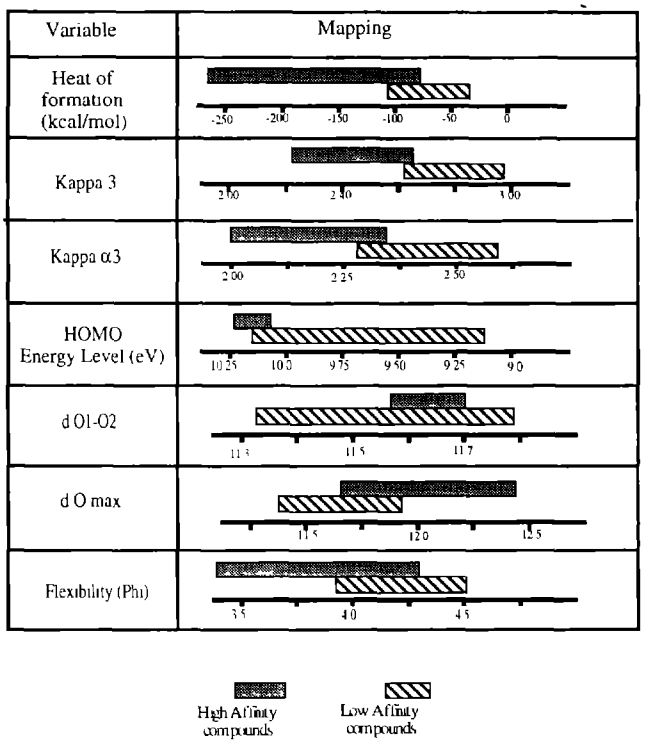


Fig 4. Variable mapping of MR affinity.

Electronic constraints. As presented in figure 4, a narrow domain of HOMO energy (or first ionisation potential values) and a low heat of formation are involved in high affinity towards the hMR.

Structural constraints. As previously described in similar studies a distance between two electronegative centers, distance O_1–O_2, or the maximal interdistance between two oxygens ($d_{O_{max}}$) constitute an important feature of affinity. The range of distances that are favourable for high affinity is narrow for the distance O_1–O_2 (around 11.60 Å) as well as for the $d_{O_{max}}$ (> 11.6 Å).

All these distributions have been validated by significant cluster analysis studies (scores in tables IV and V).

The total score of discrimination between high and low affinity compounds deduced from a set involving two constraints (distance O_1–O_2 and HOMO energy level or heat of formation and distance O_1–O_2) is 100%. The totality of high-affinity molecules is situated at the score maximum of constraints respect, all the other molecules with lower affinity obtain lower scores.

Table IV. Cluster significance analysis, MR affinity constraints. Number of random trials: 100 000; population high affinity: 8; low affinity: 5; standardisation used: mean/standard deviation.

Variable used	*Class*	*Significance (95% range estimate)*
HOMO energy level (eV)	High affinity	0.988 (± 0.002)
	Low affinity	0.356 (± 0.009)
Heat of formation (kcal/mol)	High affinity	0.952 (± 0.004)
	Low affinity	0.991 (± 0.002)
Distance O_1–O_2	High affinity	0.996 (± 0.001)
	Low affinity	0.003 (± 0.001)
$d_{O_{max}}$	High affinity	0.933 (± 0.005)
	Low affinity	0.980 (± 0.003)
Kappa 3 (Kier)	High affinity	0.975 (± 0.003)
	Low affinity	0.878 (± 0.006)
Kappa α3	High affinity	0.796 (± 0.008)
	Low affinity	0.815 (± 0.008)
Flexibility (Phi)	High affinity	0.542 (± 0.01)
	Low affinity	0.606 (± 0.01)

Table V. Cluster significance analysis, MR affinity. Number of random trials: 100 000; population high affinity: 8; low affinity. 5; standardisation used: mean/standard deviation. The score of discrimination for all these sets of constraints = 100%.

Set of constraints used	*Class*	*Significance (95% range estimate)*
Distance O_1–O_2 and heat of formation	High affinity	0.999 (± 7 x 10^{-4})
	Low affinity	0.248 (± 0.008)
HOMO and distance O_1–O_2	High affinity	0.999 (± 5 x 10^{-4})
	Low affinity	0.061 (± 0.005)
Distance O_1–O_2 Kappa 3	High affinity	0.992 (± 0.002)
	Low affinity	0.235 (± 0 008)
HOMO Distance O_1–O_2 Heat of formation	High affinity	0.999 (± 5 x 10^{-4})
	Low affinity	0.287 (± 0.008)

These variables constitute the minimal set to be used for a theoretical screening (in silico screening) of the high-affinity molecules. Other associations of two or more constraints can be used and show an equivalent efficacity.

This very simple method performs, in this particular case, much better than discriminant analysis especially from an explicative point of view.

Stepwise discriminant analysis

Taking into account the dual classification previously described, a discriminant analysis was undertaken. Table VI represents the best results as a function of the selected variables by the stepwise discriminant analysis.

These results are in good agreement with the constraints deduced from the variable mapping study.

Analysis of the activity towards the human mineralocorticoid receptor

In order to specify the structural and physico-chemical requirements involved in the agonist–antagonist activities of our set of compounds, a dual classifications was taken into account, dividing the molecules into two groups:

- the group of agonist molecules: Aldosterone, 18OVP, DOC, 21DP, Corticosterone, Cortisol, Dexamethasone (agonist group);
- the group of antagonist molecules: Progesterone, 18VP, Triamcinolone, 18EP, 18DCP1, 18DCP2 (antagonist group).

Stepwise discriminant analysis

Taking into account the dual classification previously described, a stepwise discriminant analysis was undertaken. Table VII represents the best result. As is shown by the selection of the Carbo Gaussian lipophilicity index, using Aldosterone as a reference, the agonist molecules of this study have a spatial distribution of their molecular lipophilic potential similar to those induced by the natural agonist (aldosterone).

Variable mapping

As previously described in other work [44], the intervals of variation of the structural, lipophilic and electronic properties of agonist compounds are narrower than those of the antagonists. In this study nearly all the property distributions expressed as constraints obey this rule (fig 5 and table VIII). In particular all the physico-chemical properties of the agonist compounds are very close to those of the natural agonist Aldosterone.

A set of three constraints involving the Log *P*, the distance O_1–O_2 and the RY component of the ellipsoidal volume allows a 100% discrimination between agonists and antagonists. This study points out that the lipophilicity of ligands plays an important role in the determination of their agonist character. The design of agonists of the hMR must be carried out within the respect of strict lipophilic contigencies, as well for their lipophilic potential distribution than for their global lipophilicity expressed by Log *P*.

Conclusion

The use of variable mapping combined with the use of cluster significance analysis of the molecular characteristics, responsible for the biological activity, has allowed us to extract some precise structural and physico-chemical molecular characteristics necessary

Table VI. Stepwise discriminant analysis results, MR affinity.

Item	*Variable*	*% well-predicted*
Topological indices	Kappa 3	93%
	Kappa α3	100%
3D Autocorrelation vectors (step = 1 Å)	Bin 12	76%
	Bin 8	92%
Structural properties	$d_{O_{max}}$	85%
	I_1	85%
	3D AV Bin 8	100%
Electronic and structural properties	Formation heat	85%
	HOMO energy	85%
	$d_{O_1\text{-}O_2}$	100%
Electronic and lipophilic properties	Formation heat	85%
	Log *P*	100%

Table VII. Stepwise discriminant analysis results, MR activity (agonist–antagonist).

Variable	*ASP carbo lipophilicity*	*O_2 charge*	*$d_{O_1\text{-}O_2}$*
% well predicted	77%	93%	93%

Table VIII. Cluster significance analysis, activity. Number of random trials: 100 000; population agonist: 7; antagonists: 6; standardisation used: mean/standard deviation.

Set of constraints used	*Class*	*Significance (95% range estimate)*
Distance O_1–O_2	Agonist	0.997 ($\pm 4 \times 10^{-4}$)
Log *P*, RY	Antagonist	0.096 ($\pm$ 0.002)

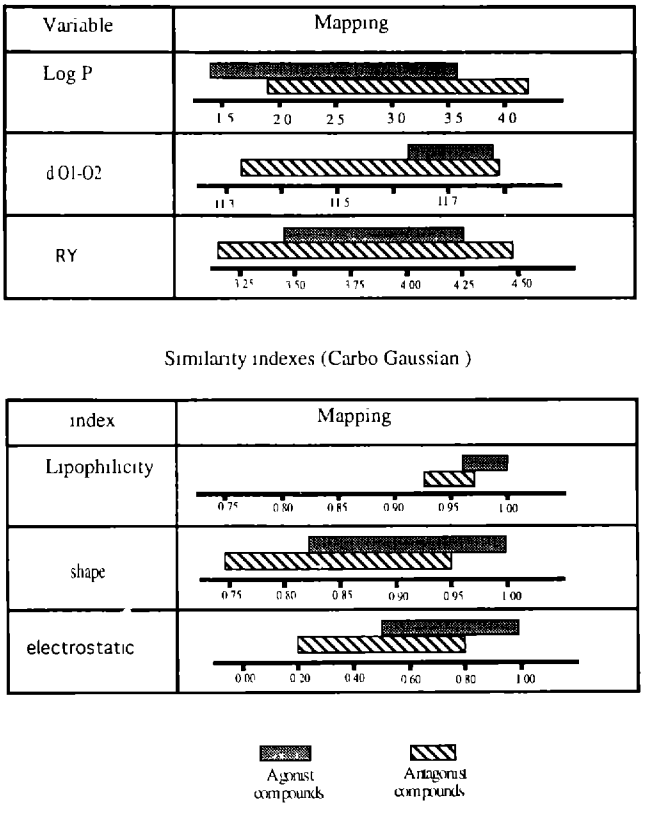


Fig 5. Variable mapping of MR activity.

for the rational prediction of the agonist or antagonist compounds of the hMR possessing high affinity for this receptor. The model presented in this paper constitutes a theoretical screening that can be used for rational drug design of anti-mineralocorticoids. Parallel, work on the modelling of the receptor site is currently in progress in order to localise the criteria of interaction that we have revealed here.

Acknowledgments

This study was supported by a grant from the CNRS (ACC SV5). The authors aknowledge a grant of computer time from the CNUSC (France). Thanks are due to C Royer from the University of Wisconsin–Madison for helpful comments.

References

1 Arriza JL, Weinberger C, Cerelli G, Glaser TM, Handelin B, Housman E, Evans RM (1987) *Science* 237, 268–275
2 Evans RM (1988) *Science* 240, 889–895
3 Tsai MJ, O'Malley BW (1994) *Annu Rev Biochem* 63, 451–486
4 Rafestin-Oblin ME, Couette B, Radanyi C, Lombes M, Baulieu EE (1989) *J Biol Chem* 264, 9304
5 Schulman G, Daniel V, Cooper M, Alnemri ES, Maksymowych AB, Litwack G (1992) *Receptor* 2, 181–194
6 Caamano CA, Morano MI, Patel PD, Watson SJ, Akil H (1993) *Biochemistry* 32, 8589–8595
7 Nemoto T, Ohara-Nemoto Y, Sato N, Ota M (1993) *J Biochem* 113, 769–775
8 Lombès M, Binart N, Oblin ME, Joulin V, Baulieu EE (1993) *Biochem J* 292, 577–583
9 Trapp T, Rupprecht R, Castren M, Reul JM, Holsboer F (1994) *Neuron* 13, 1457–1462
10 Trapp T, Holsboer F (1995) *Biochem Biophys Res Com* 215, 286–291
11 Couette B, Fagart J, Jalaguier S, Lombès M, Souque A, Rafestin-Oblin ME (1996) *Biochem J* 315, 421–427
12 Corvol P, Claire M, Oblin ME, Geering K, Rossier BC (1981) *Kidney Int* 20, 1–6
13 Wambach G, Casals-Stenzel J (1984) Structure–activity relationship of spironolactone derivates. In *Adrenal steroid antagonism*, De Gruyter, Berlin, New York, 291–313
14 Sutanto W, de Kloet ER (1991) *Med Res Rev* 11, 617–631
15 Claire M, Faraj H, Grassy G, Aumelas A, Rondot A, Auzou G (1993) *J Med Chem* 36, 2404–2407
16 Landau RL, Bergenstal DM, Lugibihl K, Kascht ME (1955) *J Clin Endocrinol Metab* 15, 1194–1215
17 Wambach G, Higgins JR (1978) *Endocrinology* 102, 1686–1693
18 Rafestin-Oblin ME, Couette B, Barlet-Bas C, Cheval L, Viger A, Doucet A (1991) *Am J Physiol* 260, 828–832
19 Rupprecht R, Reul JM, Van Steensel B, Spengler D, Soder M, Berning B, Holsboer F, Damm K (1993) *European J Pharmacol* 247, 145–154
20 Viger A, Coustal S, Perard S, Marquet A (1988) *Tetrahedron* 44, 1127–1134
21 Viger A, Coustal S, Perard S, Chappe B, Marquet A (1988) *J Steroid Biochem* 30, 469–472
22 Coustal S, Fagart J, Davioud E, Marquet A (1995) *Tetrahedron* 51, 3559–3570
23 Souque A, Fagart J, Couette B, Davioud E, Sobrio F, Marquet A, Rafestin-Oblin ME (1995) *Endocrinology* 136, 5651–5658
24 Davioud E, Fagart J, Souque A, Rafestin-Oblin ME, Marquet A, *J Med Chem* (in press)
25 Claire M, Rafestin-Oblin ME, Michaud A, Corvol P, Venot A, Roth-Meyer C, Boisvieux JF, Mallet A (1978) *FEBS Lett* 88, 295–299
26 Binart N, Lombès M, Rafestin-Oblin ME, Baulieu EE (1991) *Proc Natl Acad Sci USA* 88, 10681–10685
27 Graham FL, Van der Eb AJ (1973) *Virology* 52, 456–467
28 De Wet JR, Wood KV, Deluca M, Helsinki DR, Subramani S (1987) *Mol Cell Biol* 7, 725–737
29 Herbomel P, Bourachot B, Yaniv M (1984) *Cell* 39, 653–662
30 MAD® 2.20, Oxford Molecular Ltd, Magdalen Centre, Oxford Science Park, Sandford-on-Thames, Oxford OX4 4GA, UK
31 Stewart JJP, MOPAC 6.0, QCPE # 455
32 a) Meyer AY, Richards WG (1991) *J Computer Aided Mol Design* 5, 427
b) ASP® 3.02, Oxford Molecular Ltd, Magdalen Centre, Oxford Science Park, Sandford-on-Thames, Oxford OX4 4GA, UK
33 Burt C, Richards WG, Huxley P (1990) *J Comp Chem* 11, 1139–1146
34 TSAR® V2.31, Oxford Molecular Ltd, Magdalen Centre, Oxford Science Park, Sandford-on-Thames, Oxford OX4 4GA, UK
35 Balaban AT (1982) *Chem Phys Lett* 89, 399
36 Wiener H (1947) *J Am Chem Soc* 69, 2636
37 Randic M (1975) *J Am Chem Soc* 97, 6609
38 Hall LH, Kier LB (1992) *Review in Computational Chemistry* (Lipkowitz KB, Boyd DB, ed), 367
39 Broto P, Moreau G, Vandycke C (1984) *Eur J Med Chem* 19, 61–84
40 a) Ghose AK, Pritchett A, Crippen GM (1988) *J Comput Chem* 7, 565–577
b) Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) *J Chem Inf Comput Sci* 29, 163
41 Manly BFJ (1986) *Multivariate Statistical Methods. A Primer*, Chapman and Hall
42 McFarland JW, Gans DJ (1986) *J Med Chem* 29, 505–514
43 McFarland JW, Gans DJ (1987) *J Med Chem* 30, 46–49
44 Grassy G, Trape P, Bompart J, Calas B, Auzou G (1995) *J Mol Graph* 13, 356–367