Molecular Recognition in Purinergic Receptors. 1. A Comprehensive Computational Study of the *h*-P2Y$_1$-Receptor†

Dan T. Major and Bilha Fischer*

Department of Chemistry, Gonda-Goldschmied Medical Research Center, Bar-Ilan University, Ramat-Gan 52900, Israel

Received March 24, 2004

P2Y receptors (P2Y-Rs) are attractive pharmaceutical targets due to their involvement in the modulation of many tissues and organs. The lack of experimental structural data on P2Y-Rs impedes structure-based drug design. The need to elucidate the receptor's molecular recognition, together with the limitations of previous receptor models, triggered the construction of a new molecular model for the *h*-P2Y$_1$-R. Therefore, a *h*-P2Y$_1$-R model was constructed by homology modeling using the 2.6 Å crystal structure of bovine rhodopsin as a template and subsequently refined by constrained molecular dynamics (MD) simulations in a fully hydrated lipid bilayer environment. ATP was docked into the receptor binding site, followed by binding site refinement using Monte Carlo and MD simulations. Analysis of the *h*-P2Y$_1$-R−ATP complex suggests that the triphosphate moiety is tightly bound by a multitude of interactions possibly including a Mg^{2+} ion, the ribose ring is not involved in specific interactions, and the adenine ring is bound via N1, N7, and N^6. The molecular recognition of the *h*-P2Y$_1$-R was further probed by ATP derivatives modified on the adenine ring, and correlated with EC$_{50}$ values for these derivatives. Analysis of receptor:ligand complexes and quantum mechanical studies on model compounds support the role of both steric and electronic effects in improving H-bonding (via N1 and N^6) and π-stacking interactions. The computed *h*-P2Y$_1$-R model was validated with respect to our previous biochemical results. We believe that this new model of the *h*-P2Y$_1$-R provides the means for understanding phenomena such as the ligand's potency and receptor subtype selectivity.

Introduction

The P2Y-receptors (P2Y-Rs) comprise a family of G-protein coupled receptors (GPCRs) activated by extracellular ATP, **1**. These ubiquitous receptors have been studied extensively, employing various approaches from molecular biology to pharmacology and medicinal chemistry.[1−6] The P2Y-receptors are attractive pharmaceutical targets due to their involvement in the modulation of many tissues and organs under both normal and pathophysiological conditions.[7−11] Several potential drugs targeting the P2Y-Rs have been developed by applying structure−activity relationship (SAR) considerations.[12−18] However, rational structure-based drug design requires prior knowledge of the structure of the target receptor and the ligand's principal binding site. This essential information necessitates the availability of the pure P2Y-receptor subtype in sufficient amounts for further structural analysis by X-ray crystallography or NMR studies. Although we have successfully isolated the P2Y$_1$-R on a novel affinity chromatography column,[19] many obstacles remain, such as the difficult crystallization of membrane receptors and their complex NMR analysis. These obstacles preclude one from obtaining structural data for P2Y-receptors.

The need to identify the ligand's binding site triggered alternative studies, mainly site-directed mutagenesis and homology modeling, based on a related receptor template. A vast and comprehensive work was performed with that respect on the P2Y$_1$-R by Jacobson and co-workers.[20] The first model, of chick P2Y$_1$-R, was constructed on the basis of the density map of bovine rhodopsin (*b*-Rh) and was verified by the incorporation of site-directed mutagenesis and ligand affinity data.[21,22] Integration of the available data enabled the identification of several amino acid residues in transmembrane helices (TMs) 3, 5, 6, and 7 that are responsible for ligand recognition. Later, Moro et al. introduced the "cross docking" procedure to obtain energetically refined 3D structures of ligand-*h*-P2Y$_1$-R complexes, thus simulating the ligand-induced reorganization of the receptor.[23] This model was used to identify residues that are responsible for binding P2Y$_1$-R agonists and antagonists. Moro et al. suggested that Arg128, Lys280, and Arg310 are important for recognition of the ATP phosphate moiety, while Gln307 and Ser314 create the binding pocket for the adenine ring. In addition to the previously defined ATP-binding transmembrane cleft, Moro also proposed meta binding sites, defined by the extracellular loops EL2 and EL3.[24,25] The observed ligand potency was considered to be affected by binding at both principal and lower energy meta-binding sites. A ligand's path of access was also proposed, from the extracellular side to the principal binding site. The conformational requirements of various ligands were evaluated by this model, and a set of conformationally restricted antagonists revealed that sugar N conformation is favored by the P2Y$_1$-R.[26] The crystal structure of *b*-Rh obtained at 2.8 Å resolution[27,28] enabled Jacobson et al. to propose an improved model for the *h*-P2Y$_1$-

† This work has been supported by Innodia Inc, Montreal, Canada.
* Corresponding author. Fax: 972-3-5351250. Tel: 972-3-5318303. E-mail: bfischer@mail.biu.ac.il.

10.1021/jm049772m CCC: $27.50
Published on Web 08/03/2004

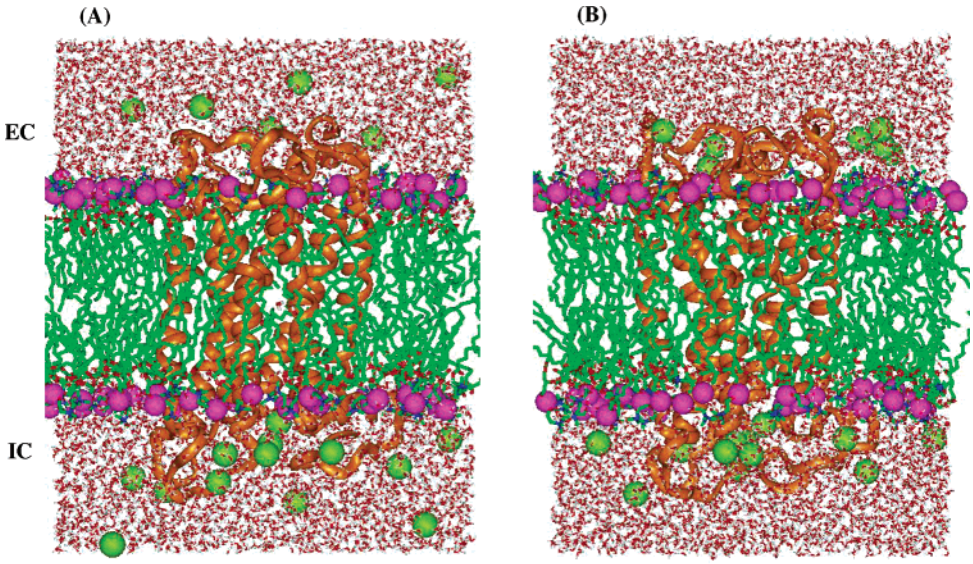


Figure 1. *h*-P2Y$_1$-R embedded in a fully hydrated lipid-bilayer environment: (A) initial system configuration and (B) system after 700 ps.

R.[29] The refined model was similar to the previous one with respect to the side chains composing the binding-site. However, this model differed in the length of the helices and their mutual orientation within the helical bundle. Another difference between the models was due to the location of EL2, which was deeply embedded in the TM region, as in rhodopsin. In this model, the only loop included was EL2.

The above-mentioned molecular models of the P2Y$_1$-R enabled the understanding of agonists and antagonists recognition and helped the design of more potent and subtype-selective P2Y-R ligand-based drugs. However, all of these models were truncated forms of the P2Y$_1$-R, excluding the intracellular loops (ILs). Furthermore, the models were not generated in their natural environment, i.e., in a fully hydrated lipid bilayer, but rather optimized in the gas-phase. These limitations, together with the recent availability of a bovine-rhodopsin crystal structure at a higher resolution than previously used,[28] justify the continued search for a refined P2Y$_1$-R model.

In addition to the construction of an improved *h*-P2Y$_1$-R model, we present here a comprehensive structural analysis of the new model, with a view to elucidate the mode of molecular recognition. This analysis includes investigation of TM helical kinks, interhelical interactions stabilizing the helical bundle, and the structure and possible function of intra- and extracellular loops. The molecular recognition of ATP (**1**) and ADP (**2**) by the *h*-P2Y$_1$-R is also discussed, on the basis of the new model. The stereoelectronic recognition of various P2Y$_1$-R agonists, **3**−**6**, modified at the adenine ring, is analyzed on the basis of docking studies and on molecular properties of these agonists as calculated by quantum mechanical methods.[30,31] Furthermore, in the companion paper we discuss the origin of the diastereoselectivity of the P2Y$_1$-R based on the current receptor model and chiral nucleotide probes.[32]

Results

The *h*-P2Y$_1$-R is a phospholipase C-activating ATP receptor present in a variety of tissues and organs.[1,12,20] The molecular weight of the *h*-P2Y$_1$-R is 42 071 Da and

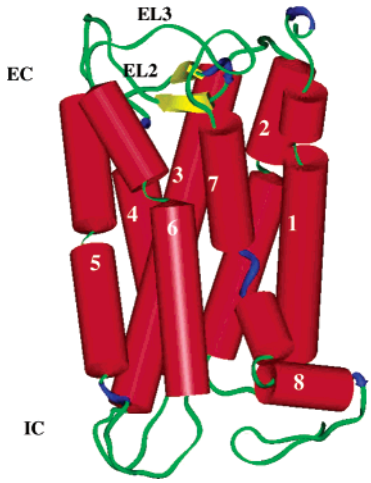


Figure 2. General topology of the *h*-P2Y$_1$-R: red, α-helix; yellow, *β*-sheet; blue, turn; green, random coil..

it has 373 amino acids. The identity of *h*-P2Y$_1$-R with other mammalian P2Y$_1$-Rs is high: bovine, 96%; rat, 95%; mouse, 94%; chick, 85%; turkey, 85%. Yet, the identity of *h*-P2Y$_1$-R with other *h*-P2Y-R subtypes is modest: *h*-P2Y$_2$-R, 30%; *h*-P2Y$_4$-R, 33%; *h*-P2Y$_6$-R, 32%; *h*-P2Y$_{11}$-R, 27%; *h*-P2Y$_{12}$-R, 22%; *h*-P2Y$_{13}$-R, 22%. Bovine rhodopsin (*b*-Rh), a G-protein coupled receptor, shares only 19% identity with the *h*-P2Y$_1$-R. Despite this low sequence homology, the crystal structure of *b*-Rh was chosen as a template for the homology modeling of *h*-P2Y$_1$-R, since both share the same functionality as GPCRs and it is well established that the topology is conserved better than the sequence.[33]

The molecular modeling of the *h*-P2Y$_1$-R involved creating an initial guess structure for the receptor by homology modeling using the 2.6 Å X-ray crystal structure *b*-Rh,[28] in its inactive state, as a template. The initial guess structure was subsequently refined in a fully hydrated lipid bilayer environment with constrained MD simulations as described in the Methods section. The final model includes the *h*-P2Y$_1$-R, lipid bilayer, bulk water, and counterions (Figure 1).

Receptor Structure Analysis. The refined receptor model is similar in structure to the initial homology

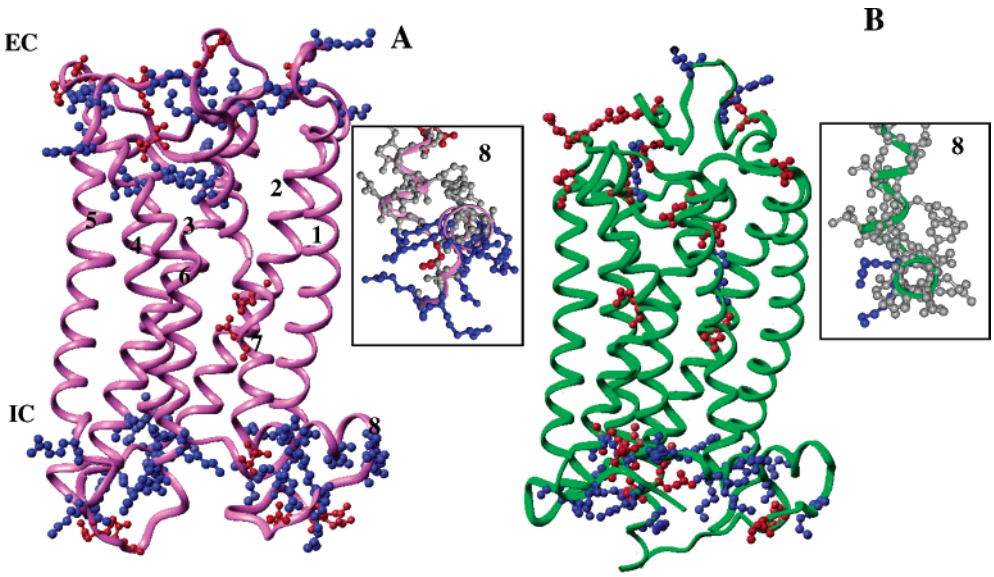


Figure 3. Comparison of the distribution of charged amino acid residues of the *h*-P2Y$_1$-R (A) and *b*-Rh (B): red, Asp/Glu; blue, Arg/Lys. Insets: top-view of cytoplasmic helix (8).

model. The root-mean-square difference (rmsd) between the initial guess structure and the final refined model is 1.9 Å when accounting for all heavy atoms. To assess the extent of change in each structural motif, we superimposed the initial and final receptor models using the Cα trace of the TM region as a reference. This yielded a Cα atom rmsd of 0.75 Å for the TM region, 2.45 Å for the ELs, and 2.13 Å for the ILs. To assess the structural change of each TM helix individually, we also superimposed each helix individually. This yielded Cα atom rmsd (Å) for TM1−7: 0.64, 0.79, 0.66, 0.64, 0.47, 0.54, and 0.61.

The TM Region Topology. A basic assumption regarding the *h*-P2Y$_1$-R homology model is that the helices making up the TM bundle are tilted relative to one another in a manner similar to *b*-Rh. However, the specific topology of the *h*-P2Y$_1$-R is unique and is determined by its amino acid sequence (Figure 2).

In the *h*-P2Y$_1$-R there are a total of 14 Pro residues; seven of these are located in the TM helical region. With the exception of TM3, all TM helices have a Pro residue. These Pro residues stabilize significant kinks in most of the *h*-P2Y$_1$-R TM helices (Figure 2). The greatest kinks are seen for TM helices 6 and 7, while TM3 is entirely straight. The Pro residues in TM4, -5, -6, and -7 are highly conserved in the GPCR family. Interestingly, the conserved prolines are located in the helices that form the binding site in *b*-Rh [27] and the presumed *h*-P2Y$_1$-R binding site.[23] On the other hand, the Pro residues in TM1 and -2 are unique to the *h*-P2Y$_1$-R.

Gly residues, which introduce conformational flexibility, are also abundant in the *h*-P2Y$_1$-R. The *h*-P2Y$_1$-R has a total of 19 Gly residues of which eight are within the TM region: two in each of TM1, -3, and -5 and one in each of TM2 and -7. In TM1 and -5, the Gly residues are part of a GXXG motif that has been shown to provide significant flexibility in TM helices.[34] Another six Gly residues are located at the ends of the TM helices. These Gly residues probably play a role in adding flexibility to the helix−loop interface region.

Interactions in the TM Region. The stability of the *h*-P2Y$_1$-R within the lipid bilayer may be ascribed to several correlated physical and chemical factors.

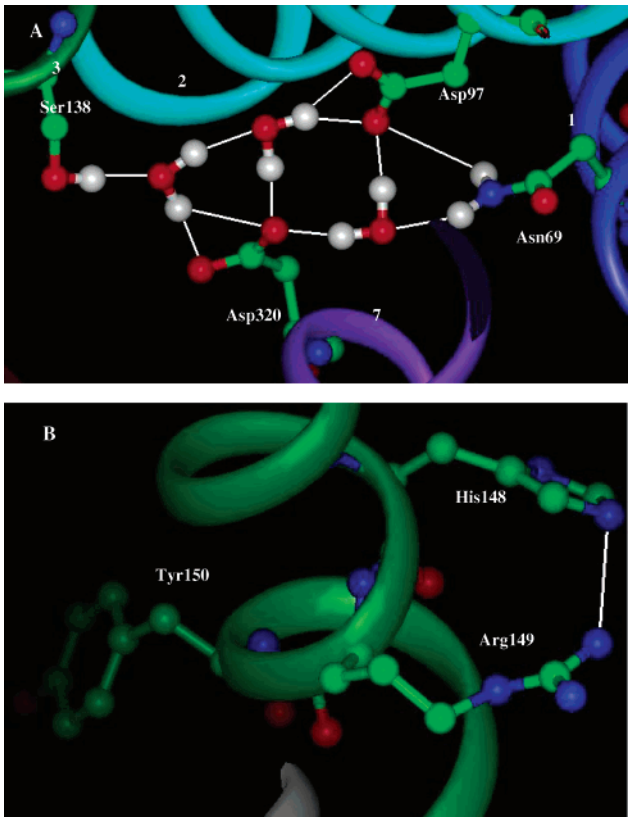


Figure 4. Snapshot of conserved GPCR residues in the *h*-P2Y$_1$-R after 500 ps of MD simulation in a lipid bilayer: (A) polar interactions between TM helices: Asn69 (TM1), Asp97 (TM2), and Asp320 (TM7); (B) polar interactions within TM3, His148 and Arg149.

First, the overall organization of the *h*-P2Y$_1$-R within the cell membrane is typical for GPCRs with seven antiparallel TM helices. The calculated dipole moments of these helices were found to be large, ranging from 100 to 300 D. This leads to a strong electrostatic interaction between the antiparallel helices' dipoles, thus stabilizing the helical bundle. Second, the TM bundle is stabilized by numerous hydrophobic interactions between hydrophobic amino acids in the TM region and the membrane

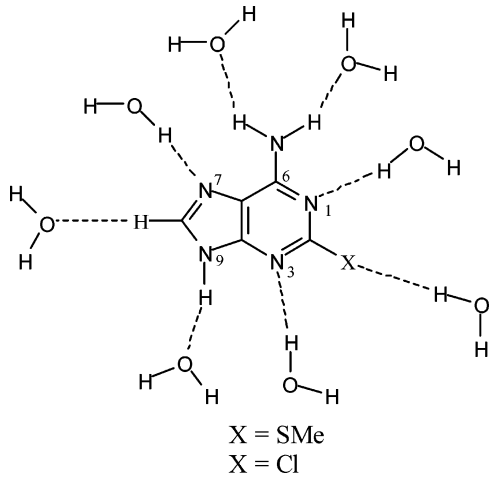


Figure 5. Interaction geometries of water−adenine derivatives

lipid alkyl chains. Third, the *h*-P2Y$_1$-R possesses many polar and charged residues at the lipid−water interface, creating a polar ring at the intra- and extracellular faces (Figure 3A). These residues anchor the receptor within the membrane via many electrostatic and polar interactions with the polar lipid headgroups. Amino acids involved in these interactions include Trp, Tyr, Lys, and Arg residues. All these strong interactions enhance the stability of the helical bundle in the membrane environment.

In addition, the TM helical bundle in *h*-P2Y$_1$-R is stabilized by several interhelical polar and hydrophobic interactions. A cluster of interactions between helices 1, 2, 3, 6, and 7 is found approximately midway through the TM region. This cluster involves several conserved GPCR residues such as Asn69 (TM1), Asp97 (TM2), and Asp320 (TM7) (Figure 4A). Asp320 is part of the highly conserved (N/D)PXXY motif.[27] In the current model, these interactions are partly bridged by three water molecules, as in *b*-Rh.[28] Additional clusters of interhelical interactions include Tyr324 (TM7) and Ser87 (TM2) interacting with each other through a bridging water molecule and the conserved Asn92 (TM2) and Trp176 (TM4) residues interacting through H-bonds.

A well-known motif in GPCRs is the (D/E)R(Y/W) motif in TM3. Mutational analysis reveals that the aspartate/glutamate-arginine ionic interaction plays a pivotal role in receptor activation.[35] In the *h*-P2Y$_1$-R, this triad consists of His148, Arg149, and Tyr150 in TM3. Thus, the unprotonated His148 nitrogen may interact with the neighboring guanidium moiety in Arg149 (Figure 4B).

The large number of conserved GPCR residues in the P2Y$_1$-R that are thought to be involved in interhelical interactions support the assumption that the general topology of the *h*-P2Y$_1$-R is similar to that of *b*-Rh.

Loop Regions. The *h*-P2Y$_1$-R displays a multitude of positively charged residues in the extracellular and intracellular loops as well as in the C-terminal (Figure 3A). Many positively charged amino acid residues in the ELs and at the ends of the TM helices are apparently unique to the *h*-P2Y$_1$-R (compare with *b*-Rh, Figure 3B). Together, the positively charged residues in the ELs may play a role in guiding the negatively charged nucleotide into the binding site, as suggested by Moro et al.[24]

Table 1. Comparison of Ab Initio and Empirical Interaction Energies (kcal/mol) and Geometries (Å) between Water and Adenine (Ade),[a] 2-MeS-Ade (2MeS), and 2-Cl-Ade (2Cl)

	ab initio (HF/6-31G*) × 1.16		CHARMM	
interaction	R_{min}	E_{min}	R_{min}	E_{min}
Ade N1−HW	2.10	−6.99	1.89	−6.97
Ade H2−OW	2.49	−1.51	2.45	−1.55
Ade N3−HW	2.12	−7.09	1.90	−7.07
Ade H61−OW	2.00	−5.30	1.85	−5.35
Ade H62−OW	2.08	−4.66	1.89	−4.54
Ade N7−HW	2.08	−7.09	1.89	−7.16
Ade H8−OW	2.39	−2.95	2.37	−3.27
Ade H9−OW	2.01	−7.25	1.84	−7.25
2MeS N1−HW	2.14	−7.28	1.91	−7.80
2MeS S2−HW	3.00	−1.48	2.60	−1.24
2MeS N3−HW	2.14	−7.45	1.90	−7.83
2MeS H61−OW	2.05	−4.89	1.88	−4.47
2MeS H62−OW	2.00	−4.67	1.85	−4.20
2MeS N7−HW	2.08	−6.87	1.89	−6.72
2MeS H8−OW	2.35	−3.72	2.36	−3.36
2MeS H9−OW	1.98	−8.06	1.83	−7.64
2Cl N1−HW	2.14	−6.96	1.91	−7.08
2Cl Cl2−HW	3.53	−0.43	2.87	−0.44
2Cl N3−HW	2.15	−6.93	1.91	−7.10
2Cl H61−OW	2.03	−5.98	1.85	−5.76
2Cl H62−OW	1.97	−6.15	1.82	−5.94
2Cl N7−HW	2.11	−6.18	1.90	−6.67
2Cl H8−OW	2.33	−3.92	2.35	−3.66
2Cl H9−OW	1.96	−8.62	1.83	−7.31

[a] From ref 65.

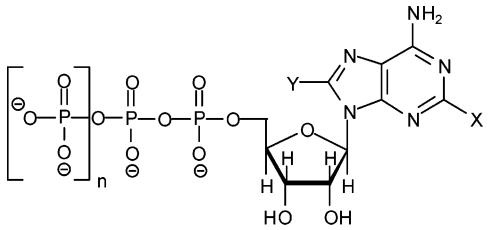
Our MD simulations support such an electrostatic directing effect. For instance, the chloride counterions used in the simulations were initially placed at random positions in the extra- and intracellular side of the triphasic box (Figure 1A). In the course of the simulation, the negatively charged chloride ions diffuse toward the loops to maximize the electrostatic interactions (Figure 1B).

Interestingly, the small intracellular helix (H8) at the C terminal has a clear amphiphilic character (Figure 3 inset). The side of H8 exposed to the membrane is hydrophobic, while the side exposed to the cytoplasm is highly charged with many Arg and Lys residues (Figure 3A). It has been suggested that such a GPCR extracellular helix may be involved in interaction with the G-protein.[36] Mutational analysis of some of these residues may elucidate their importance in signal transduction from the *h*-P2Y$_1$-R to the G-protein. In comparison, in *b*-Rh the intracellular helix carries a much lesser charge (Figure 3B).

Validation of the Computed *h*-P2Y$_1$-R Model. To validate the computed model of the *h*-P2Y$_1$-R we searched for a correlation of the binding interaction of the receptor with ligands **3**−**6** (Scheme 1) with EC_{50} values for these ligands.[19,37] For the computational studies of the interaction of ATP derivatives **3**−**6** with the *h*-P2Y$_1$-R, the CHARMM force field was extended to include new parameters. These parameters account for MeS, BuS, and Cl substitutions at the C2-position and BuS substitution at the C8-position of the adenine ring in ATP.

Adenine Base Modifications. The partial atomic charges of the adenine ring were fitted to ab initio interaction energies with individual water molecules (Figure 5). Thus, the interaction energies between a water molecule and the different H-bonding positions on the adenine base were calculated. The interaction

Scheme 1. ATP and Derivatives Docked into *h*-P2Y$_1$-R



1. X = H, Y = H, n = 1

2. X = H, Y = H, n = 0

3. X = MeS, Y = H, n = 1

4. X = BuS, Y = H, n = 1

5. X = Cl, Y = H, n = 1

6. X = H, Y = BuS, n = 1

energies obtained using HF/6-31G* and the new CHARMM parameters are presented in Table 1. The values for adenine are presented for comparison. As the ab initio results indicate, no dramatic change in interaction distances or energies occurs as a result of substitution with thiomethyl or chloro groups at the C2-position. However, the 2-MeS substitution lowers the interaction energy at the N1- and H8-positions by −0.3 and −0.8 kcal/mol, respectively, relative to adenine. On the other hand, the interaction energy at the N^6-position is increased by 0.4 kcal/mol. The 2-Cl-substitution lowers the binding energy at the N^6-position by −1.5 and −0.7 kcal/mol relative to adenine. The interaction energy with the H8-hydrogen is likewise lowered by −1.0 kcal/mol. Interaction energy at the N7-position is increased by 0.9 kcal/mol. These effects of electron-donating (ED) and electron-withdrawing (EW) substituents are similar to what was observed in our previous works.[30,31] The CHARMM interaction distances and energies agree with the ab initio results. The correlation coefficients, R^2, from a least-squares fit between ab initio and empirical distances and energies for 2-Cl-adenine are 0.92 and 0.95, respectively. In the case of 2-MeS-adenine, the correlation coefficients are 0.89 and 0.98 for interaction distances and energies, respectively. Thus, the agreement between the ab initio and empirical results is satisfactory.

Docking Studies. To investigate the molecular recognition of the adenine moiety in ATP by the *h*-P2Y$_1$-R, we apply a combined computational approach. This approach involves docking and MD studies of molecular probes **1**−**6** with the *h*-P2Y$_1$-R and quantum mechanical calculations of the corresponding reduced adenine models.

The probes we selected for the analysis of stereoelectronic recognition determinants of the *h*-P2Y$_1$-R are based on our previous studies[30,37] and include ATP analogues substituted at C2 or C8 with an electron-donating group (thioether) or at C2 with an electron-withdrawing group (Cl).

Analysis of the *h*-P2Y$_1$-R Binding-Site and ATP Recognition Mode. The ATP binding site identified in the current *h*-P2Y$_1$-R model is in accordance with the earlier work of Jacobson and co-workers.[21,23,29] The *h*-P2Y$_1$-R forms a narrow crevice approximately 5 Å below the lipid-bilayer plane that fits an extended

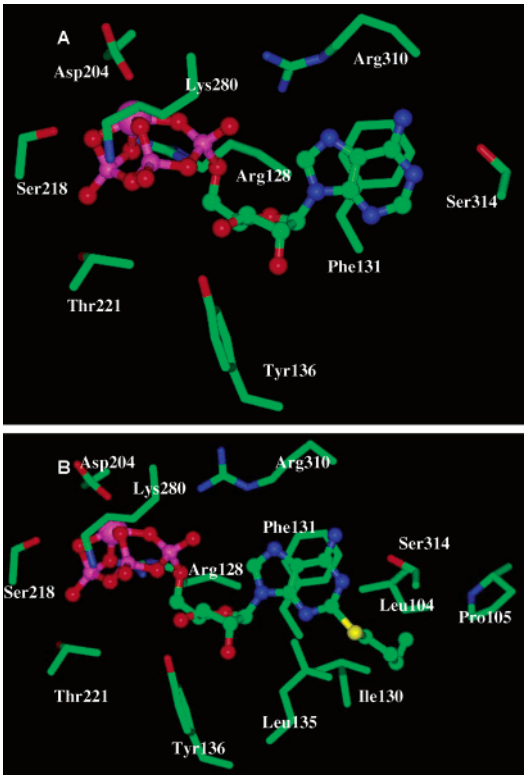
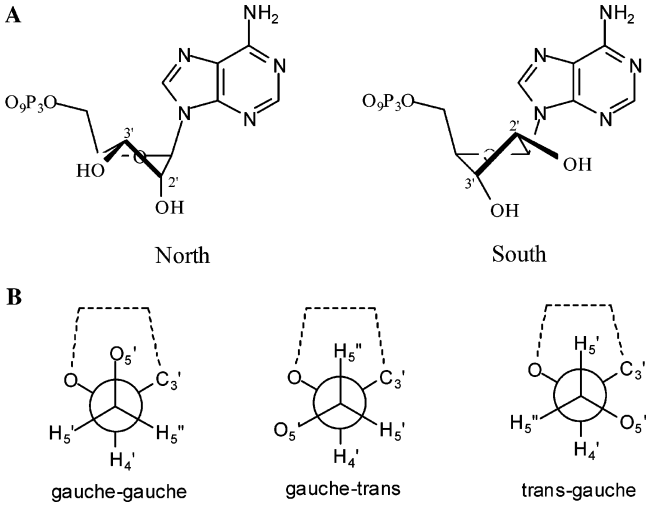


Figure 6. MD snapshots of *h*-P2Y$_1$-R-adenine nucleotide complexes: (A) ATP bound to the *h*-P2Y$_1$-R and (B) 2-BuS-ATP bound to the *h*-P2Y$_1$-R.

Scheme 2. Ribose and Exo-Cyclic C4′−C5′ Conformations in ATP



conformation of ATP (Figure 6). All TM helices participate in forming the binding site, although only TM3, -5, -6, and -7 are involved in specific interactions. In the current model, EL2 apparently plays an important part of the binding pocket, as is the case for *b*-Rh.[27] The β-sheet of EL2 is deeply embedded in the helical bundle and may function as a lid on the binding site, together with EL3 (Figure 2).

Moro et al. proposed that the phosphate moiety in ATP is coordinated by Arg128 (P_α and P_β), Tyr136 (P_α), Thr222 (P_γ), Lys280 (P_α), and Arg310 (P_β).[23] Subsequently, Kim et al. suggested that the phosphate chain is located adjacent to the assumed β-sheet of EL2.[29] According to this suggestion, the phosphate moiety is

Table 2. Average Interaction Distances (Å) between Heavy Atoms of Selected Residues of the *h*-P2Y$_1$-R and Adenine Nucleotides during 100 ps of MD Simulations in a Fully Hydrated Lipid Bilayer[a]

	ATP		ADP		2-Cl-ATP		2-MeS-ATP		2-BuS-ATP	
N1···Ser314	3.09	±0.28	3.39	±0.46	2.96	±0.22	2.99	±0.17	3.06	±0.18
N6···Ser314	3.38	±0.47	4.02	±0.76	3.33	±0.31	3.22	±0.35	3.04	±0.16
N6···Arg310(O)	2.80	±0.12	3.00	±0.30	2.91	±0.17	2.85	±0.13	2.88	±0.14
N7···Arg310(NE)	3.91	±0.47	4.38	±0.54	3.44	±0.26	3.58	±0.41	3.66	±0.43
N7···Arg310(NH1)	3.58	±0.56	4.22	±1.23	2.97	±0.13	3.59	±0.50	5.00	±0.46
O3···Tyr96	3.03	±0.31	3.50	±0.26	2.95	±0.24	2.92	±0.19	2.88	±0.21
P_α···Arg128(NH1)	3.52	±0.26	3.63	±0.28	3.14	±0.19	3.23	±0.25	3.37	±0.30
P_α···Arg128(NH2)	4.20	±0.23	3.25	±0.39	4.38	±0.23	4.27	±0.32	4.73	±0.40
$P_{\beta/\gamma}$···Arg128(NH1)[b]	2.98	±0.20	5.46	±0.37	2.79	±0.17	2.91	±0.33	2.73	±0.20
$P_{\beta/\gamma}$···Arg128(NH2)[b]	2.71	±0.11	4.98	±0.20	2.85	±0.19	2.97	±0.31	2.91	±0.22
$P_{\beta/\gamma}$···Ser218[b]	2.77	±0.19	2.71	±0.15	2.64	±0.09	2.79	±0.19	2.74	±0.16
$P_{\beta/\gamma}$···Thr221[b]	2.77	±0.16	2.88	±0.25	2.72	±0.12	2.80	±0.21	3.59	±0.92
P_β···Lys240	2.68	±0.09	2.61	±0.07	2.70	±0.09	2.70	±0.10	2.69	±0.09
P_α···Arg310(NH1)	3.80	±0.28	5.45	±1.19	4.00	±0.29	3.81	±0.22	6.47	±0.25
P_α···Arg310(NH2)	2.73	±0.11	3.85	±0.90	2.76	±0.14	2.80	±0.17	3.87	±0.27
P_α···Mg^{2+}	1.87	±0.06	2.10	±0.26	1.89	±0.06	1.89	±0.06	1.87	±0.05
P_β···Mg^{2+}	1.85	±0.05	1.85	±0.05	1.84	±0.05	1.97	±0.09	1.85	±0.05
$P_{\beta/\gamma}$···Mg^{2+} [b]	1.87	±0.06	1.82	±0.04	1.85	±0.05	1.89	±0.06	1.86	±0.07
	1.87	±0.05	3.74	±0.06	1.88	±0.06	1.90	±0.09	1.86	±0.06

[a] Interaction distances are measured between heavy atoms. Fluctuations in the distances are included to the right of the distance. [b] P_β for ADP and P_γ for ATP and derivatives.

loosely bound in a sterically unconstrained pocket and P_γ is not involved in specific interactions. However, this is counterintuitive, considering the strong interactions that the highly charged phosphate chain is expected to form. A reason for this phosphate binding mode seems to be a need to avoid repulsive interaction with Asp204 that is positioned facing the binding pocket.[29]

Furthermore, Moro et al. suggested that the ribose ring interacts with His132, His277, and Ser317, while adenine interacts with Gln307, Arg310, and Ser314.[23] Subsequent studies indicated that the ribose moiety serves mainly as a linker unit, while Gln307 is probably not directly involved in agonist recognition.[29]

A striking feature in the current model is the large number of strong interactions involved in ATP binding (Table 2, Figure 6). The main interactions involving Arg128, Lys280, Arg310, and Ser314 agree with previous mutational analysis and docking studies,[22,23,29] although differences exist.

The Triphosphate Moiety Is Tightly Bound by a Multitude of Interactions with Charged and Polar Residues and Possibly a Mg^{2+} Ion. The ATP phosphate chain is involved in a multitude of interactions with residues in TM3, -5, -6, and -7 (Table 2, Figure 6). The α-phosphate interacts with Arg128 and Arg310, while the *β*-phosphate interacts with Lys280. The *γ*-phosphate moiety interacts directly with Arg128, Ser218, and Thr221. Additional interactions of the triphosphate group via bridging water molecules are possible with His132, Tyr136, Thr222, and His277. For instance, Tyr136 and His277 may each interact with the phosphate group via a water molecule. Such water molecules play an important role in the recognition of the ribose ring in adenine nucleotides,[38] and one might also expect a similar role in the recognition of the phosphate and adenine moieties. His132 and Thr222 seem to interact strongly with Lys280, holding it firmly in place. The Arg128, Lys280, and Arg310 residues were found to be essential for receptor activity by mutational analysis experiments, while the remaining residues are assumed to play a modular role in ligand binding and receptor activation.[22,23] Particularly, Thr221 is likely to be involved in terminal phosphate recognition and was found to be more important than Thr222 for receptor activity.

To the best of our knowledge, no experimental evidence has been provided for the presence of a metal ion in the *h*-P2Y$_1$-R binding-site. However, we hypothesize, on the basis of our MD simulations and the chiral discrimination described in the companion paper,[32] that the ATP phosphate chain coordinates with a Mg^{2+} ion, which in turn is coordinated with Asp204 (Figure 6). Asp204 was found to be an important amino acid residue by mutational analysis, and it was suggested to play a yet unknown role in the physiological activation of the *h*-P2Y$_1$-R.[39]

An appealing argument in favor of the presence of a Mg^{2+} ion is that this ion leads to a fully neutralized binding site. Thus, the negative charge of Asp204 and the four negative charges of ATP[40] may be neutralized by Arg128, Lys280, Arg310, and a divalent metal ion. Furthermore, Asp204 is conserved in several other P2Y-Rs (P2Y$_{2,4,6}$-R), indicating a common function. On the basis of both these arguments and the results cited in our companion paper,[32] we introduced Mg^{2+} ion in our simulations of the *h*-P2Y$_1$-R-ATP complex. An intriguing question follows the possible presence of a Mg^{2+} ion in the *h*-P2Y$_1$-R binding site: How many and which of the phosphate groups coordinate with the Mg^{2+} ion? The current simulations support the coordination of Mg^{2+} ion with the α, *β*, *γ* phosphates on the basis of the topology of the binding site (Figure 6, Table 2). Yet, further experiments are needed to clarify this point.

Nonetheless, it is important to stress that MD simulations performed with ATP in the absence of Mg^{2+} ion (results not shown) yielded very similar molecular recognition for the phosphate moiety. The main changes are related to the conformation of the phosphate chain, to allow for metal ion coordination. The nature of the interactions with the receptor remains largely the same. An important difference is that Asp204 moves out of the binding pocket and may form ionic interactions with Arg128 or Lys280.

The Ribose Ring Is Not Involved in Specific Interactions with the *h*-P2Y$_1$-R. Moro and Jacobson

suggested that the ribose ring plays a role mainly as a linker unit between the adenine ring and the phosphate chain.[39] This role is supported in this study, where few specific interactions with the ribose ring are observed. The O2′ group is not involved in any interactions, while O3′ may interact directly with Tyr136 (Table 2) or via a water molecule with His277.

The Adenine Ring Is Loosely Bound by H-Bonds and π-Stacking. The adenine moiety is loosely bound, as reflected in the fluctuations in the distances between the receptor and the adenine moiety (Table 2), interacting with amino acid residues from TM7 (Figure 6). This is in contrast to the phosphate binding, which is tight and rigid. Ser314 interacts via H-bonds to N1 and N^6, while the backbone carboxyl group of Arg310 interacts with N^6. Additionally, the Arg310 side chain interacts with N7. Such interactions between the adenine moiety and a positively charged residue were recently found to be a general adenine base recognition motif.[41] Interaction between N^6 and Tyr58 is also possible via a bridging water molecule. In the current model, the adenine moiety is also involved in π-stacking with Phe131 with a distance ranging from 3.7 to 4.0 Å. Phe131 was found to play a modular role in *h*-$P2Y_1$-R activation.[22] The adenine ring is also in proximity to Phe276.

The results for phosphate, ribose, and adenine recognition are consistent with mutational analysis data and previous modeling studies.[22,23]

ADP Recognition Mode. The mode of ADP recognition by *h*-$P2Y_1$-R is very similar to that of ATP, the main differences being the phosphate moiety, as expected (Table 2). The interactions between P_β and Arg128 and between P_α and Arg310 are weaker than in the case of ATP. Moreover, both the adenine and ribose rings are more loosely bound than in the case of ATP.

Validating the Receptor Model and Probing Stereoelectronic Requirements of the *h*-$P2Y_1$-R. Our previous studies demonstrated that ATP derivatives substituted with a thioether group at the C2-position make potent *h*-$P2Y_1$-R agonists, while the same substitutions at the C8-position yield inactive ligands.[37] Thus, to validate the receptor model, we docked both C2- and C8-substituted thioether ATP derivatives (**3**, **4**, and **6**) into the model.

2-BuS-ATP binds to the receptor in a mode similar to that of ATP (Table 2) with additional hydrophobic interactions via the thiobutyl group. In the current model the thioether group fits in a hydrophobic pocket comprised of Leu104, Pro105, Ile130, Val133, and Leu135. Val61 and Val101 could also possibly be involved. Docking of 2-MeS-ATP resulted in similar results, although with fewer hydrophobic interactions. In the MD simulations of the C2-thioether derivatives, the average fluctuation in interaction distances for the adenine moiety is lower than in the case of ATP, indicating a more rigid fit (Table 2). On the other hand, 8-BuS-ATP could not fit into the receptor binding pocket without disrupting supposed crucial interactions. 8-BuS-ATP was docked first in a syn conformation, based on our conformational analysis of this molecule.[37] In this conformation interactions of the adenine moiety with Arg310 and Ser314 were lost. Likewise, when 8-BuS-ATP was docked in an anti conformation, interactions of both the adenine ring and the phosphate chain with Arg310 were lost. Both Arg310 and Ser314 are known to be crucial to receptor activity.[22,23] Indeed, in a previous study we showed a drastic reduction of 8-BuS-ATP potency, as compared to ATP (EC_{50} ca. 10 μM).[19,37]

Similar to the 2-thioalkyl ATP derivatives, 2-Cl-ATP also yields a more rigid fit with the receptor, similar to the 2-thioether ATP analogues (Table 2). This improved steric fit may partially account for its improved affinity for the receptor.[37]

Modeling the π-Stacking Interactions between the Adenine Ring and Phe131. To explore the potential of adenine, 2-MeS-adenine, and 2-Cl-adenine to participate in π-stacking, their adducts with benzene were analyzed by quantum mechanical calculations. In all cases, a tight interaction is observed with short intermolecular distances. The π-stacking interaction distance with benzene is shorter both in the case of 2-MeS-adenine and 2-Cl-adenine (3.39 Å) than in the parent base (3.44 Å). This stronger interaction is also reflected in the MP2/cc-pVDZ interaction energies. The π-stacking energy with benzene was computed as −3.76, −4.92, and −4.71 kcal/mol for adenine, 2-MeS-adenine, and 2-Cl-adenine, respectively.

It is worth noting that the π-stacking energy is sensitive to both basis set and the relative orientation of the interacting π-systems.[42] In a study on the benzene dimer, it was found that the interaction energy increased by a factor of approximately three when enlarging the basis set from cc-pVDZ to cc-pVQZ. Moreover, it was found that the slipped-parallel benzene dimer is considerably more stable than the parallel dimer. Thus, the current results, which used a small basis set and considered only the parallel orientation, should be considered a low estimate for the real π-stacking energy in these systems.

$P2Y_1$-R Induces a Unique Conformation of Bound ATP. The conformational flexibility of ATP plays a crucial role in its molecular recognition by proteins. The protein binding site induces a unique conformation of the relatively flexible ATP, which may be different from that of free ATP in solution.

Information regarding the conformation of free ATP has been obtained from X-ray crystallography, spectroscopy, and theoretical calculations.[43] The relative orientation of the adenine base and the ribose moiety is determined by the glycosidic angle, χ, and the predominant conformation in solution is the anti-conformation. The ribose ring in ATP may adopt a variety of conformations that may be described by the so-called "pseudorotation cycle". The dominant conformations of nucleosides and nucleotides in aqueous solution, in DNA and RNA, and when bound to proteins are the Northern (2′-exo, 3′-endo) and Southern (2′-endo, 3′-exo) conformations (Scheme 2A).[43] Moreover, the conformation around the exocyclic C4′−C5′ bond determines the relative orientation of the ribose ring and the phosphate chain (Scheme 2B).

Our computational analysis of the receptor−ligand complex included also conformational aspects. For this purpose we compared the conformational properties of ATP in the current MD simulations with the conformations of protein-bound adenine nucleotides in experimentally solved structures.

Table 3. Average Dihedral Angles (deg) for Selected Adenine Nucleotides Complexed with the *h*-P2Y$_1$-R during 100 ps of MD Simulations in a Fully Hydrated lipid bilayer[a]

	ATP		ADP		2-Cl-ATP		2-MeS-ATP		2-BuS-ATP	
C4−N9−C1−O4	239.0	±10.9	218.0	±15.3	248.3	±10.7	241.9	±11.7	244.2	±9.2
O4−C4−C5−O5	−64.4	±9.3	−63.4	±8.0	−75.3	±7.9	−67.9	±6.8	−64.4	±6.7
ribose pseudorotation angle	43.5	±4.9	45.9	±4.2	41.5	±4.9	45.0	±4.1	43.8	±3.9
ribose puckering	29.3	±9.1	27.3	±7.9	31.7	±9.8	27.0	±6.6	29.6	±6.1

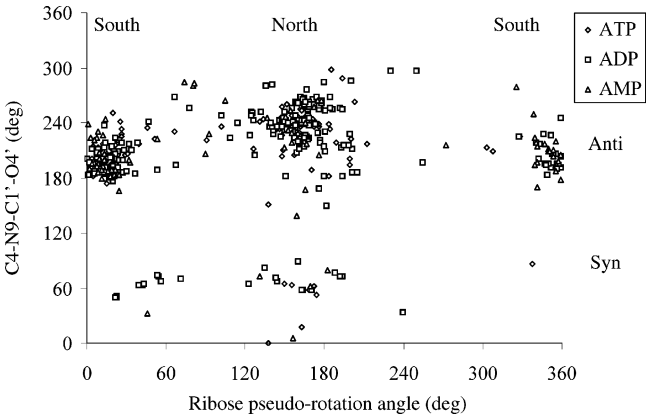


Figure 7. Correlation plot of conformations of protein-bound adenine nucleotides. The pseudorotation ribose angle is plotted against the glycosidic dihedral angle, C4−N9−C1′−O4′.

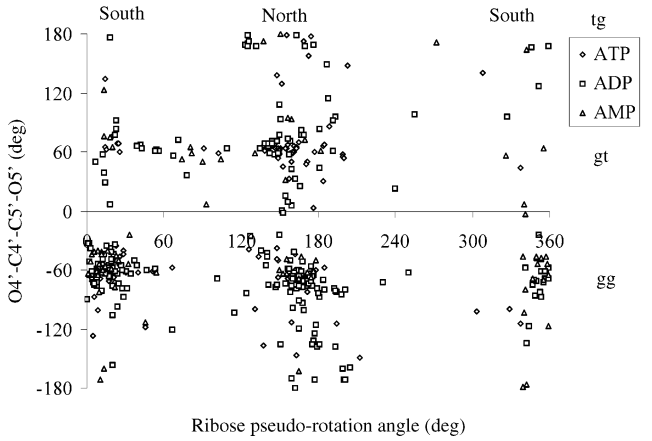


Figure 8. Correlation plot of conformations of protein-bound adenine nucleotides. The pseudorotation ribose angle is plotted against the exocyclic dihedral angle O4′−C4′−C5′−O5′.

To the best of our knowledge, no conformational analysis of all currently available protein bound adenine nucleotides has been performed. Therefore, we present in Figures 7 and 8 the conformational analysis of protein-bound ATP, ADP, and AMP from the Protein Data Bank. The most probable conformations of protein-bound adenine nucleotides are then compared with the conformations of ATP obtained during the simulations of the *h*-P2Y$_1$-R−ATP complex in a fully hydrated lipid bilayer environment.

From the data set of protein-bound adenine nucleotides, a clear preference for the anti conformation around the glycosidic dihedral angle is observed and the sugar adopts either the Northern or Southern conformations (Figure 7). In the few cases where the syn conformation is observed, the sugar conformation shows a slight preference for the Southern conformation (Figure 7). The average sugar puckering value is 36.7 ± 10.4 (Not shown in the figures).

In the current *h*-P2Y$_1$-R−ATP complex simulations, ATP is in the anti conformation (Table 3). It is likely that Arg310 supports this conformation by interaction with both the N7 of the adenine ring and the α-phosphate moiety. Over the course of 200 ps of MD simulation of the complex, the average χ-value is 239.0° with a fluctuation of ±10.9°. The ribose ring is in the Northern conformation with a pseudorotation value of 43.5 ± 4.9°, and the ring puckering is 29.3 ± 9.1. Indeed, the preferred sugar conformation at the P2Y$_1$-R has been found to be the Northern conformation. This finding is based on studies employing synthetic nucleotides, which are locked in either the Northern or Southern conformations.[44]

From the data set of protein-bound nucleotides, the conformation around the exocyclic C4′-C5′ bond shows a clear preference for the *gg* rotamer (Figure 8). Namely, the phosphate chain is oriented above the ribose ring. The *gt* rotamer is also frequently seen, while the *tg* conformation is scarcely populated. Moreover, when the ribose is in the Northern conformation, the *gg* conformation is preferred (Figure 8), probably due to a weak hydrogen bond between H8 of the adenine ring and O5′ of the sugar. It is interesting that this preference for the *gg* conformation, which is observed in aqueous solution,[43] is also seen in protein-bound adenine nucleotides. In the current brief MD simulation of ATP bound to the *h*-P2Y$_1$-R, the *gg* conformation is exclusively occupied with a value of −64.4° ± 9.3°.

The *h*-P2Y$_1$-R-bound C2-substituted ATP derivatives all have similar conformational properties to those of receptor-bound ATP. This emphasizes the similar binding mode of these ligands and ATP. This is in contrast to the 8-BuS-ATP derivative, which is likely in the syn conformation,[37] thereby interrupting important interactions with the receptor and reducing drastically the affinity to the *h*-P2Y$_1$-R.

Discussion

The Stereoelectronic Character of the Substituted Adenine May Affect Molecular Recognition of the ATP Derivative by the *h*-P2Y$_1$-R. A crucial question regarding the molecular recognition of ATP derivatives by the *h*-P2Y$_1$-R is the role of stereoelectronic effects. Both 2-MeS-ATP and 2-Cl-ATP are much more potent agonists than ATP, with EC$_{50}$ values of 8, 72, and 2800 nM, respectively.[45] Thus, the *h*-P2Y$_1$-R is sensitive to changes on the adenine moiety. Moreover, the *h*-P2Y$_1$-R apparently favors substitutions at the C2-position of ATP, regardless of the electronic nature of the substituent. On the other hand, substitutions at the C8-position yield inactive ligands.[19,37] This may indicate that a stereo effect is present, but not an electronic one. However, on the basis of our docking studies combined with quantum mechanical calculations on model compounds we propose that the 2-MeS- and 2-Cl-substitutions may improve interactions with the receptor due to both steric effects and electronic effects. According

to the current *h*-P2Y$_1$-R model, the adenine moiety occupies approximately the same position for ATP, 2-MeS-ATP, 2-BuS-ATP, and 2-Cl-ATP. In the docked complexes of 2-MeS-ATP, 2-BuS-ATP, and 2-Cl-ATP, the C2-substituent fits into a hydrophobic pocket of the *h*-P2Y$_1$-R (Figure 6B). This additional hydrophobic interaction due to the C2-substituent is likely to increase the affinity of the ligand for the receptor. Moreover, the C2 substituent yields a more rigid steric fit between the *h*-P2Y$_1$-R and the ligand, thereby reducing the loss in entropy upon binding, e.g. the ligand is preoriented in the correct binding mode.

The current *h*-P2Y$_1$-R model indicates that recognition of the adenine moiety is via the N1-, N^6-, and N7-positions of the base (Table 2). The nature of the recognition of the adenine moiety was further investigated by substituting electron-donating (ED) and electron-withdrawing (EW) groups at the adenine ring. These groups induce electronic effects that are not readily addressed by MD simulations, both due to the approximate nature of the force fields and the complex potential energy surface. Therefore, our previous studies[30,31] employed quantum mechanics calculations to elucidate the electronic effects due to various ED and EW substitutions at the C2-position on the adenine ring. These studies showed that ED substituents, such as thiomethyl, increase the electron density at the N1-position in adenine, thus increasing its potency as H-bonding acceptor. The same substituents had minimal effect on the N^6-position. On the other hand, EW substituents, such as chlorine, decrease electron density on the N^6-position, thereby increasing the potential of the N^6 hydrogens to participate in H-bonding as donors. These earlier findings are supported here by the current quantum mechanical calculation of interaction energies between individual water molecules and adenine, 2-MeS-adenine, and 2-Cl-adenine (Figure 5 and Table 1). Indeed, interaction between the N1-position of the adenine moiety and a water molecule is more favorable in the case of 2-MeS-adenine, than in the case of adenine. Likewise, the interaction between the N^6-position in 2-Cl-adenine and a water molecule is more favorable than in adenine. Thus, it seems that 2-MeS-ATP and 2-Cl-ATP gain improved interactions via the N1- and N^6-positions, respectively. On the other hand, both 2-MeS-ATP and 2-Cl-ATP have weaker interactions at the N7-position than adenine. When summing up the individual contributions from interactions via N1, N^6, and N7 and accounting for the fact that both sulfur and chlorine participate in weak H-bonding interactions with the protein via Ser314, 2-MeS-ATP and 2-Cl-ATP have more favorable interactions by a total of 1.2 and 1.7 kcal/mol, respectively (Table 1). Moreover, the current QM level of theory does not account for dispersion effects that may be present for sulfur and chlorine.

π-Stacking Interactions May Play an Important Role in Stabilizing the ATP-Receptor Complex. In the current *h*-P2Y$_1$-R model, the adenine moiety is involved in π-stacking interactions with Phe131. The importance of π-stacking interactions for the recognition of substituted ATP analogues was explored here by QM calculations on model compounds. We performed model π-stacking calculations on benzene and adenine, 2-MeS-adenine, and 2-Cl-adenine by employing a second-order perturbation theory QM method. These model calculations showed that the π-stacking energy may be substantial (low estimate ca. -5 kcal/mol) and that the adenine moieties in both 2-MeS-ATP and 2-Cl-ATP are expected to interact more favorably with Phe131 than the adenine ring in ATP. The reason for this electronic effect could be that in the 2-MeS-derivative the adenine ring functions as a charge *donor* molecule in a π-stacking charge-transfer complex, while the adenine base in the 2-Cl-derivative plays the role of a charge *acceptor*.

Thus, the modeling studies indicate a similar binding mode for the various ATP derivatives and suggest that steric effects are likely to play a crucial role in molecular recognition. Moreover, the QM calculations suggest that subtle differences in the electronic character between ATP and its C2-substituted derivatives may account for differences in affinity for the *h*-P2Y$_1$-R.

Conclusions

A model for the *h*-P2Y$_1$-R was constructed by homology modeling and subsequently refined in a fully hydrated lipid bilayer. The model was validated by comparing the molecular recognition by the *h*-P2Y$_1$-R, predicted by docking studies, with available mutational and biochemical data.

As indicated by our modeling results, the *h*-P2Y$_1$-R is stabilized by numerous interactions, including dipolar interactions of antiparallel TM helices, hydrophobic and polar interactions with the lipid bilayer, and intrahelical interactions.

The proposed primary binding site in the receptor is located approximately 5 Å below the lipid bilayer plane, within TM3, -5, -6, and -7 and below EL2. There, ATP, in an extended conformation, is tightly held in place by numerous direct and indirect (via bridging water molecules) interactions of the triphosphate moiety with the receptor, including H-bonds and electrostatic interactions with Lys and Arg residues. In addition, we hypothesize that the involvement of Mg^{2+} ion, possibly coordinated with the conserved Asp204, completes the tight network by fixing the triphosphate moiety in the receptor. This is supported by accumulating evidence described in our companion paper.[32]

In addition to the strong recognition network observed for the triphosphate moiety, another important network of interaction, although significantly weaker, is observed for the adenine ring. Despite the lesser contribution of the adenine moiety to the stabilization of the ATP$-$receptor complex, such interactions are extremely important, since they determine the receptor-subtype selectivity. The current model suggests nonspecific base recognition, due to π-stacking interactions between the adenine moiety and Phe131. Specific H-bonding interactions to N1, N^6, and N7 are provided by Arg310, Ser314, and possibly Tyr58. These interactions are seemingly enhanced in the presence of electron donating and withdrawing substituents at the C2-position, due to electronic effects. Moreover, the C2-substituents yield a more rigid fit between the adenine moiety and the receptor. In the case of 2-MeS- and 2-BuS-ATP, hydrophobic interactions are likely to enhance the interaction with the receptor. On the other hand, substitutions at the C8-position are not tolerated by the receptor, due to disruption of important interactions. This confirms

the role of both steric and electronic effects of the adenine moiety in improving affinity to the *h*-P2Y$_1$-R.

Both the flexibility of the adenine binding mode (as opposed to that of the phosphate chain) and the minute role of the ribose ring in recognition imply that chemical modifications of the adenine moiety may play a major role in modulating the affinity and receptor-subtype selectivity of a P2Y$_1$-R ligand.

We believe that our model provides the means for understanding phenomena such as a ligand's potency and receptor subtype selectivity and forms the basis for rational drug design.

Methods

Overview. The *h*-P2Y$_1$-R model was computed using bovine rhodopsin (*b*-Rh) at 2.6 Å resolution[28] as a template. A basic assumption in the current protocol is that the relative orientation of the TM helices is similar in *b*-Rh and *h*-P2Y$_1$-R, although not necessarily identical. Moreover, H-bonding and electrostatic interactions between conserved GPCR residues are assumed to be similar in the two proteins. Briefly, the protocol entails creating an initial guess structure for the receptor based on *b*-Rh. This initial model is very similar to *b*-Rh and includes all the conserved interhelical interactions in the TM region. The initial guess structure is subsequently refined in a lipid bilayer environment to improve the TM helical packing while constraining the conserved interhelical interactions and secondary structure in the TM region. The homology conserved constraints together with additional constraints on the secondary structural motifs are necessary to avoid large fluctuations in the receptor structure prior to equilibrium.

A large number of modeling studies of GPCRs have appeared recently, using the X-ray structure of *b*-Rh as a template. These studies have demonstrated that homology modeling may yield useful models despite the low sequence identity between *b*-Rh and other GPCRs.[46] In the current study, the intial model obtained from homology modeling results in a binding site that is too narrow to fit the endogenic ligand ATP, in agreement with previous findings.[47] Therefore, a model refinement is necessary. It has been found that minimization or MD simulation of homology models may actually reduce the quality of homology models.[48] However, the effect of minimization and MD simulation on the quality of homology models is likely to depend on the kind of protein modeled (soluble vs membrane protein) and the surrounding environment employed in the computations (gas phase, aqueous solution, or lipid bilayer). It is clear that unphysical gas-phase refinements will introduce errors, as has been noted.[49] However, it was demonstrated that constrained MD simulations in explicit solvent yield improved homology models.[49] In the particular case of GPCRs, it has been demonstrated that MD refinement in a lipid environment may yield useful models.[50] Moreover, the hydrated lipid environment plays a role in the ligand's recognition. Specifically, any energy calculation on a charged molecule like ATP requires treatment of charge screening by water molecules that are expected to be present in the binding site. Additionally, in the case of ATP derivative **4**, the long alkyl chain is expected to be in close vicinity to the lipid hydrophobic tails. Thus, accounting for the hydrated bilayer environment is necessary both for model construction and molecular recognition studies.

The current model was validated with respect to biochemical data. ADP, ATP, and adenine-modified ATP derivatives **3**−**6** were docked into the proposed binding site and the respective binding modes were correlated with our EC$_{50}$ data for these ligands.[19,37] Additionally, the amino acid residues involved in ligand binding were compared with available mutational analysis data and previous modeling studies.[22,23] In the companion paper, the quality of the model is further validated by docking several chiral ATP derivatives into the receptor model and correlating the interaction energies of the receptor−ligand complexes with the corresponding EC$_{50}$ values.[32]

The modeling protocol used here involved the following steps:

(1) homology modeling using the 2.6 Å X-ray crystal structure of *b*-Rh,[28]

(2) constrained molecular dynamics (MD) refinement of the receptor in a hydrated-lipid bilayer environment, and

(3) docking of ATP and analogues into the receptor model and refining the binding site.

Homology Modeling. Initially, multiple sequence alignment of the P2Y-Rs with known P2Y-R-like physiological effects[15] and *b*-Rh was performed using the ClustalX program.[51] Gap penalties were introduced in the secondary structure regions of *b*-Rh, namely the transmembrane (TM) regions, in the intracellular helix (H8), and in the β-sheets of the second extracellular loop (EL2). Gap penalties to the TM region are essential for correct alignment between the *h*-P2Y$_1$-R and *b*-Rh. The alignment of conserved GPCR residues in the TM regions of the P2Y-Rs was verified and manual corrections were made in the loop regions. The lengths of the secondary structure motifs were determined on the basis of *b*-Rh.

The aligned *h*-P2Y$_1$-R and *b*-Rh sequences were imported to Modeller 6,[52] where a stepwise model construction approach was adopted. The Modeller program was used because of its ability to add a variety of constraints to the energy function. Therefore, the Modeller program can be adopted to the modeling process of the receptor under investigation. Moreover, Modeller includes several energy terms from the CHARMM force field, which was used in the subsequent simulations. Thus, a smooth transition should be expected when moving from the homology modeling step to the molecular simulation step.

Two disulfide bridges were introduced: Cys42-Cys296 and Cys124-Cys202. The former is a highly conserved disulfide bridge in the GPCR family,[27] while the latter is unique to the P2Y-R family.[24,25] Constraints were added to account for differences in the proline distribution in the TM regions between *h*-P2Y$_1$-R and *b*-Rh. Thus, in helices where *b*-Rh includes Pro residues with no equivalent Pro in *h*-P2Y$_1$-R, regular helicality was imposed in the *h*-P2Y$_1$-R model. TM2 in *b*-Rh includes a π-helical turn due to two consecutive Gly residues that are part of a GGF motif. The aligned *h*-P2Y$_1$-R sequence is TLP, and therefore a π-helical turn is not expected for this receptor. Thus, regular helicality was initially assumed for TM2 and the effect of the Pro residue was accounted for in the MD simulation steps. The presence of the Pro residue imposes an initial strain on the system that slowly relaxes during the MD simulation. The remaining noncanonical features observed in the TM helices of *b*-Rh were assumed to be similar in the *h*-P2Y$_1$-R, since there is currently no reliable way to predict such anomalous helical structures based on sequence.[53] Five water molecules, involved in interhelical interactions via conserved GPCR residues, were included in the model (Termed Wat1a, Wat1b, Wat1c, Wat3, and Wat4 in the work of Okada et al. on the X-ray structure of *b*-Rh[28]). Initially, 100 models were constructed using Modeller 6, and the lowest energy model was selected for future refinement. After obtaining the initial model, loops were modeled with the loop-modeling algorithm implemented in Modeller 6.[54] Mutational studies suggest that an important ionic bridge between EL2 and EL3 may exist due to the interaction between Glu209 and Arg287.[24,25] Thus, in the loop modeling process, harmonic constraints were added between Glu209 and Arg287.

The final model contained 373 amino acid residues with heavy atoms only.

MD of the Triphasic Receptor−Lipid−Water System. Receptor−Lipid−Water System Construction. The *h*-P2Y$_1$-R model was embedded in a fully hydrated dimyristoyl phosphatidylcholine (DMPC) bilayer to account for the environmental effects on the receptor structure and on the ligand's recognition. All simulations used the CHARMM program.[55] The triphasic orthorhombic system was constructed according to a modified version of the stepwise protocol of Woolf and Roux.[56] The protocol entails the packing of hydrated lipids around the protein, full hydration of the intra- and extracel-

lular regions, followed by stepwise MD equilibration of the system, and, finally, a MD production run.

To reduce the size of the simulation, the receptor was truncated at the N- and C-termini. Thus, the model receptor spans residue 40−323. This was done to reduce the large number of water molecules needed to solvate the flexible terminal regions. This is not expected to influence the packing of the TM helices or the description of the principal binding site, which is assumed to be in the TM region. Moreover, the N-terminal includes several possible glycosylation sites, although the nature of these sugars is not known. The presence of several sugars is expected to influence the conformation of the N-terminal and modeling of this region would be highly speculative. Thus, the current model cannot address questions regarding the role of the N-terminal in early stages of ligand binding. Hydrogens were added using the HBUILD procedure in CHARMM. The Arg, Lys, Asp, and Glu amino acid residues were charged, whereas His was neutral. The intrinsic low resolution of a homology model does not allow for a meaningful assignment of His protonation sites. The total charge of the truncated receptor is +20. The microscopic system consists of the *h*-P2Y$_1$-R (316 residues), 81 DMPC lipids (40 in the top layer and 41 in the bottom layer), 5353 water molecules, and 20 Cl^- counterions, totaling 30 842 atoms. The chloride ions were added to the water on the intra- and extracellular side. This constitutes the central unit of the system with dimensions of approximately 55 × 64 × 83 Å^3. The total charge of the system is zero. The membrane normal is oriented along the *z*-axis, and the center of the bilayer is located at $z = 0$. Periodic boundary conditions were applied in all directions.

Initially, the question of the orientation of the receptor model relative to the membrane layer normal vector was addressed. The cryoelectron microscopy data for frog rhodopsin (*f*-Rh) of Unger and Schertler et al.[57] yields the angle of each TM helix relative to the membrane normal. Thus, the orientation of *f*-Rh was used as a template to rigidly orient the *h*-P2Y$_1$-R model relative to the lipid bilayer normal. Initially, the protein was aligned along the *z*-axis and translated in the *z*-direction so that the center of the TM helical region was located at the origin. An exhaustive conformational search was performed by rigidly rotating the receptor model around the *z*-axis and comparing the rmsd between the *h*-P2Y$_1$-R TM angles and the TM angles of *f*-Rh.

To create a representative initial structure as close to equilibrium as possible, the lipids were taken from a preequilibrated and prehydrated library of lipid molecules.[58] To determine the initial position of the lipid molecules, the lipid polar heads were represented by large vdW spheres with a cross-sectional area of 64 Å^2. This is approximately the average cross-sectional area of a single DMPC molecule.[59] The upper layer was placed at $z = 17$ Å, while the bottom layer was set at $z = -17$ Å. The position of the vdW spheres was obtained through energy minimization and MD simulations in a periodic system with *xy*-dimensions the same as the final system. The top and bottom layers of vdW spheres were constrained to the $z = 17$ and −17 Å planes, respectively. Subsequently, the vdW spheres were replaced by hydrated lipids randomly chosen from a preequilibrated and prehydrated library of lipid molecules.[58] The lipid molecules were then rotated around the *z*-axis and translated in the *xy*-plane to minimize bad lipid−lipid and lipid−receptor contacts. Subsequently, a stepwise minimization was performed in which the vdW interactions between the lipid atoms were slowly switched on until the full vdW radii for these atoms were obtained. Thereafter, bulk water was added to the system to solvate the intra- and extracellular loops and terminal regions. Finally, the full lipid−water system was subjected to a stepwise 200 ps MD equilibration protocol.

In the course of the equilibration, many constraints were employed to allow the system to reach equilibrium smoothly, avoiding deterioration of the starting model. Initially, the receptor was kept fixed, while the lipid polar headgroups were constrained to the $z = \pm 17$ Å planes. Penetration of water through the lipid bilayer was avoided by adding an opposing force to water molecules entering the hydrophobic bilayer core. In the later stages of the equilibration, the receptor was allowed to evolve, while slowly releasing constraints. The receptor model was subjected to a set of GPCR homology constraints. These are NOE constraints between conserved amino acid residues assumed to be involved in interhelical interactions. These interactions are considered to be important for the stability of the seven-TM bundle in GPCRs.[27,28] Thus, specific interatomic contacts between conserved residues are constrained to H-bonding distances to conserve the overall structure in the TM region. Specifically, interactions between the following amino acid residues were constrained: Asn69 (TM1), Asn92 (TM2), Asp97 (TM2), Trp176 (TM4), Asn316 (TM7), and Asp320 (TM7). An additional set of constraints was added between the backbone oxygen atoms of Ser272 (TM6) and Phe276 (TM6) and Thr309 (TM7). A final set of constraints was added between Ser87 and Tyr324. Several of these interactions were bridged by water molecules as in *b*-Rh. These constraints were implemented using the NOE option of CHARMM and were enforced throughout the refinement process.

The simulations used a time step of 2 fs and all bonds involving hydrogen atoms were constrained using the SHAKE algorithm as implemented in CHARMM. A cutoff scheme of 8−12−13 Å was used for both electrostatics and vdW interactions in the MD simulations. This cutoff scheme entails turning on the switching functions at 8 Å, ignoring interactions beyond 12 Å, and keeping a list of all pair interactions within 13 Å. Both force and energy switching were used for electrostatic and vdW interactions, respectively.[60]

Binding-Site Refinement and Receptor−Ligand Interactions. The interactions between the *h*-P2Y$_1$-R and the endogenic ligand ATP were investigated using a stepwise docking protocol. This protocol involves the following steps: (1) locating the receptor binding-site and initial docking of the ligand, (2) refining the binding site side chain conformation and ligand conformation using Monte Carlo (MC) simulations, and (3) MD simulation of the receptor−ligand complex.

To identify the receptor binding site, the Connoly surface[61] of the receptor was generated, and internal clusters of cavities were identified using the SPHGEN program in DOCK 4.0.[62] The identification of the binding site was also aided by available amino acid mutational analysis data.[22,23,25] After the MD refinement of the *h*-P2Y$_1$-R there is no a priori reason to assume that the amino acid side chains in the binding site are in the correct ligand binding configuration. Thus, it is necessary to refine the binding site to allow ligand binding. Initially, ATP was docked into the receptor, and 5000 ligand configurations were generated using flexible docking with DOCK 4.0.[63] The 100 best configurations were saved for the next step.

The binding-site configuration was then refined using a MC procedure that employs the MC code of CHARMM. According to this procedure, the receptor binding site is divided into several layers surrounding the ligand. The inner layer includes all amino acid residues that are located within a layer of 1.5 Å from any ligand atom. The second layer ranges from 1.5 to 3.0 Å, while the third layer extends from 3 to 5.0 Å. Such a layering of the binding site allows for independently updating MC strategy depending on the proximity to the ligand. The frequency of MC updating was greatest for the inner layer, intermediate for the middle layer, and lowest for the outer layer. The MC moves defined for the receptor were dihedral rotations around side-chain bonds. For the inner layer, rotations of up to 180° were allowed per MC step, while for the second and third layers rotations of up to 60° and 30°, respectively, were allowed. Ligand rotations and translations of up to 30° and 0.5 Å, respectively, were allowed per MC step. The receptor was simulated for 10 000 MC steps at 310 K with each of the 100 best ligand configurations obtained in the initial docking step. The best binding site was chosen as the one with the lowest energy and in agreement with mutational analysis data.

The ligand was subsequently redocked into the refined receptor binding-site using the DOCK 4.0 program[63] with

flexible docking and thereafter refined using MC simulations. In this latter MC stage, flexibility was added to all rotatable bonds of the ligand in addition to rigid ligand rotations and translations and side-chain flexibility. This stepwise docking procedure allows for accurate placement of the ligand in the binding site using DOCK, while refinement and more reliable energy scoring is obtained using the MC simulations and the CHARMM energy function. After obtaining the final receptor−ligand configuration, binding site water molecules were added and their positions were optimized by MC simulations. The water molecules were added by placing an equilibrated water box over the binding site and keeping only the water molecules occupying vacuum points.

The final receptor−ATP complex was submitted to MD simulation in the full triphasic system. For docking analogues **2**−**6**, the ATP−*h*-P2Y$_1$-R complex was used as a template, and the new complex was refined using MC and MD simulations. The binding site was refined separately for each ligand.

Force Field Development. The CHARMM 27 force field includes parameters for ATP (**1**) and ADP (**2**).[64,65] However, to describe the various ATP derivatives introduced in the current study, new parameters must be developed. Therefore, the CHARMM 27 force field was extended to include our ATP derivatives **3**−**6**. The CHARMM energy function is described by the following expression of internal and nonbonded terms:

$$V_T = \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\phi[1 + \cos(n\phi - \delta)] + \sum_{\text{1,3pairs}} K_{UB}(S - S_0)^2 + \sum_{\text{improper}} K_{\text{imp}}(\varphi - \varphi_0)^2 + \sum_{\text{nonbonded}} \left(\epsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi D r_{ij}} \right) \quad (1)$$

In eq 1, the first five terms relate to the internal energy, while the last two terms relate to the nonbonded energy. In the current study, only parameters for the dihedral and nonbonded terms were developed, since these are the only terms that are variable in the docking procedure. K_b, K_θ, K_{UB}, and K_{imp} are the bond, angle, Urey−Bradley, and improper dihedral force constants, respectively. b, θ, S, and φ are the bond length, bond angle, Urey−Bradley 1,3 distance, and improper dihedral angle, respectively. The subscript denotes equilibrium values. The torsional term is based on a cosine function, where K_ϕ is the force constant, n is the periodicity, and δ is the phase. The nonbonded terms are described by a Lennard-Jones term plus the Coulombic term. ϵ is the Lennard-Jones well depth, $R_{\min}$ is the distance at the Lennard-Jones minimum, and q_i is the partial atomic charge. A dielectric constant of unity was employed throughout.

Adenine Base Modifications. To develop parameters for the modified-adenine base in analogues **3**−**6**, the corresponding modified bases were chosen as models that represent the target compounds. The existing vdW parameters for adenine were used for the adenine moiety in analogues **3**−**6** while for the thiomethyl group the parameters were taken from the equivalent atoms in methionine. For chlorine, various values were tested and the best results were obtained by slightly modifying the existing parameters for chloride, yielding $\sigma = 3.62$ Å and $\epsilon = 0.01$ kcal/mol.

To obtain the partial atomic charges for the adenine derivatives, the charges were fitted to reproduce ab initio QM geometries and interaction energies (scaled by 1.16)[65] with individual water molecules (Figure 5). Initially, the adenine derivatives were optimized at the ab initio QM MP2/6-31G* level,[66] while for water the TIP3 geometry[67] was used. The energy minimum distances between water and the base moiety at different geometries were determined at the QM HF/6-31G* level (Figure 5).[68] The initial force field charges were set to those of adenine, and subsequently refined until an optimal fit was obtained with the target values, i.e., the interaction distance and energy at the ab initio QM level.

The periodicity for the rotation about the C2−S exo-cyclic bond in 2-MeS- and 2-BuS-adenine was chosen to be bimodal, with two maxima on the potential energy surface (PES) for the thioalkyl group eclipsed with N1 or N3. This is in accord with the target QM PES. The dihedral force constant for the C2−S rotation was fitted to reproduce the PES obtained from ab initio quantum mechanical (QM) calculations at the MP2/6-31G* level. The rotation of the methyl groups was taken from existing parameters in CHARMM. Bond and angle terms were taken from similar existing terms in CHARMM. As approximate parameters for analogue **6** we used the parameters developed for **4**.

π-Stacking Calculations. To evaluate the potential of the adenine base derivatives to engage in π-stacking interaction, model QM calculations were performed between benzene and adenine, 2-MeS-adenine, and 2-Cl-adenine, respectively. The benzene ring and the adenine analogue rings were aligned in a parallel orientation, similar to that observed in our docking studies, and the interaction distances were optimized at the MP2/cc-pVDZ level.[66] The benzene ring was oriented in such a way that it interacted with both the pyrimidine and imidazole rings of the adenine moiety. The basis set superposition error[69] was accounted for using the counterpoise method.[70]

Acknowledgment. This work was supported in part by the BMBF-MOS, Grant No. 1812, and the Marcus Center for Medicinal Chemistry. D.T.M. acknowledges the financial support provided by the Israeli Ministry of Science.

References

(1) Bhagwat, S. S.; Williams, M. P2 purine and pyrimidine receptors: Emerging superfamilies of G-protein and ligand-gated ion channels receptors. *Eur. J. Med. Chem.* **1997**, *32*, 183−193.
(2) King, B. F.; Townsend-Nicholson, A.; Burnstock, G. Metabotropic receptors for ATP and UTP: Exploring the correspondence between native and recombinant nucleotide receptors. *Trends Pharmacol. Sci.* **1998**, *19*, 506−514.
(3) Ralevic, V.; Burnstock, G. Receptors for purines and pyrimidines. *Pharmacol. Rev.* **1998**, *60*, 413−492.
(4) Illes, P.; Klotz, K.-N.; Lohse, M. J. Signaling by extracellular nucleotides and nucleosides. *Naunyn-Schmiedberger's Arch. Pharmacol.* **2000**, *362*, 295−298.
(5) Lambrecht, G. Agonists and antagonists acting at P2X recptors: Selectivity profiles and functional implications. *Naunyn-Schmiedberger's Arch. Pharmacol.* **2000**, *362*, 340−350.
(6) Communi, D.; Janssens, R.; Suaerz-Huerta, N.; Robaye, B.; Boeynaems, J.-M. Advances in signaling by extracellular nucleotides: The role and transduction mechanisms of P2Y receptors. *Cell. Signalling* **2000**, *12*, 351−360.
(7) Boarder, M. R.; Hourani, S. M. O. The regulation of vascular function by P2 receptors: Multiple site and multiple receptors. *Trends Pharmacol. Sci.* **1998**, *19*, 99−107.
(8) Barnard, E. A.; Simon, J.; Webb, T. E. Nucleotide receptors in the nervous system- an abundant component using diverse transduction mechanisms. *Mol. Neurobiol.* **1997**, *15*, 103−129.
(9) Inoue, K. The function of ATP receptors in the hippocampus. *Pharmacol. Res.* **1998**, *38*, 323−331.
(10) Chan, C. M.; Unwin, R. J.; Burnstock, G. Potential functional roles of extracellular ATP in kidney and urinary tract. *Exp. Nephrol.* **1998**, *6*, 200−207.
(11) Jacobson, K. A.; King, B. F.; Burnstock, G. Pharmacological characterization of P2 (nucleotide) receptors. *Celltransmissions* **2000**, *16*, 3−15.
(12) Fischer, B. Therapeutic applications of ATP−(P2)-receptors agonists and antagonists. *Exp. Opin. Ther. Pat.* **1999**, *9*, 385−399.
(13) Williams, M.; Jarvis, M. F. Purinergic and pyrimidinergic receptors as potential drug targets. *Biochem. Pharmacol.* **2000**, *59*, 1173−1185.
(14) Burnstock, G.; Williams, M. P2 purinergic receptors: Modulation of cell function and therapeutic potential. *J. Pharmacol. Exp. Ther.* **2000**, *295*, 862−869.
(15) Guile, S. D.; Ince, F.; Ingall, A. H.; Kindon, N. D.; Meghani, P.; Mortimore, M. P. The Medicinal chemistry of the P2 receptor family. *Prog. Med. Chem.* **2001**, *38*, 115−187.
(16) Fischer, B.; Chulkin, A.; Boyer, J. L.; Harden, T. K.; Gendron, F.-P.; Beaudoin, A. R.; Chapal, J.; Hillaire-Buys, D.; Petit, P. 2-Thioether-5′-O−(1-thiotriphosphate) adenosine derivatives as

new insulin secretagogues acting through P2Y-receptors. *J. Med. Chem.* **1999**, *42*, 3636−3646.
(17) Fischer, B.; Yefidoff, R.; Major, D. T.; Rutman-Halili, I.; Shneyvays, V.; Zinman, T.; Jacobson, K. A.; Shainberg, A. Characterization of mini-nucleotides as P2X-receptor agonists in rat cardiocyte culture. An integrated synthetic, biochemical and theoretical study. *J. Med. Chem.* **1999**, *42*, 2685−2696.
(18) Hillaire-Buys, D.; Shahar, L. M.; Fischer, B.; Chulkin, A.; Linck, N.; Chapal, J.; Loubatieres-Mariani, M. M.; Petit, P. 2-Thioether-5′-O−(1-thiotriphosphate) adenosine derivatives as new insulin secretagogues acting through P2Y-receptors. II. Pharmacological evaluation and chemical stability of 2-benzylthioether-5′-O-(1-thiotriphosphate)-adenosine. *Drug Dev. Res.* **2001**, *53*, 33−43.
(19) Zundorf, G.; Schafer, R.; Vohringer, C.; Halbfinger, E.; Fischer, B.; Reiser, G. Novel adenosine 5′-triphosphate analogue with potential for $P2Y_1$-receptor isolation pharmacologically characterized in HEK 293 cells highly expressing the rat brain $P2Y_1$-receptor. *Biochem. Pharmacol.* **2001**, *61*, 1259−1269.
(20) Jacobson, K. A.; Jarvis, M. F.; Williams, M. Purine and pyrimidine P2 receptors as drug targets. *J. Med. Chem.* **2002**, *45*, 4057−4093.
(21) van Rhee, A. M.; Fischer, B.; van Galen, P. J. M.; Jacobson, K. A. Modelling the P_2Y purinoceptor using Rhodopsin as template. *Drug Des. Discovery* **1995**, *13*, 133−154.
(22) Jiang, Q.; Guo, D.; Lee, B. X.; van Rhee, A. M.; Kim, Y. C.; Nicholas, R. A.; Schachter, J. B.; Harden, T. K.; Jacobson, K. A. A mutational analysis of residues essential for ligand recognition at the human $P2Y_1$ receptor. *Mol. Pharmacol.* **1997**, *52*, 499−507.
(23) Moro, S.; Guo, D.; Camaioni, E.; Boyer, J. L.; Harden, T. K.; Jacobson, K. A. Human $P2Y_1$ receptor: Molecular modeling and site-directed mutagenesis as tools to identify agonist and antagonist recognition sites. *J. Med. Chem.* **1998**, *41*, 1456−1466.
(24) Moro, S.; Hoffmann, C.; Jacobson, K. A. Role of the extracellular loops of G protein-coupled receptors in ligand recognition: A molecular modeling study of the human P2Y1 receptor. *Biochemistry* **1999**, *38*, 3498−3507.
(25) Hoffmann, C.; Moro, S.; Nicholas, R. A.; Harden, T. K.; Jacobson, K. A. The role of amino acids in extracellular loops of the human P2Y(1) receptor in surface expression and activation processes. *J. Biol. Chem.* **1999**, *274*, 14639−14647.
(26) Nandanan, E.; Jang, S.-Y.; Moro, S.; Kim, H. O.; Siddiqui, M. A.; Russ, P.; Marquez, V. F.; Busson, R.; Herdewijn, P.; Harden, T. K.; Boyer, J. L.; Jacobson, K. A. Synthesis, biological activity, and molecular modeling of ribose-modified deoxyadenosine bisphosphate analogues as P2Y(1) receptor ligands. *J. Med. Chem.* **2000**, *43*, 829−842.
(27) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E. Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739−745.
(28) Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E. M.; Shichida, Y. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5982−5987.
(29) Kim, H. O.; Barak, D.; Harden, T. K.; Boyer, J. L.; Jacobson, K. A. Acyclic and cyclopropyl analogues of adenosine bisphosphate antagonists of the P2Y(1) receptor: Structure−activity relationships and receptor docking. *J. Med. Chem.* **2001**, *44*, 3092−3108.
(30) Major, D. T.; Halbfinger, E.; Fischer, B. Molecular recognition of modified adenine nucleotides by the P2Y(1)-receptor. 2. A computational approach. *J. Med. Chem.* **1999**, *42*, 5338−5347.
(31) Major, D. T.; Laxer, A.; Fischer, B. Protonation studies of modified adenine and adenine nucleotides by theoretical calculations and N-15 NMR. *J. Org. Chem.* **2002**, *67*, 790−802.
(32) Major, D. T.; Nahum, V.; Wang, Y.; Reiser, G.; Fischer, B. Molecular Recognition in Purinergic Receptors. 2. Diastereoselectivity of the *h*-$P2Y_1$-Receptor. *J. Med. Chem.* **2004**, *47*, 4405−4416 (following paper in this issue).
(33) Lesk, A. M.; Chothia, B. How different amino acid sequences determine similar protein structures: The structure and evolutionary dynamics of the globins. *J. Mol. Biol.* **1980**, *136*, 225−270.
(34) Bright, J. N.; Sansom, M. S. P. The flexing/twirling helix: Exploring the flexibility about molecular hinges formed by proline and glycine motifs in transmembrane helices. *J. Phys. Chem. B.* **2003**, *107*, 627−636.
(35) Alewijnse, A. E.; Timmerman, H.; Jacobs, E. H.; Smit, M. J.; Roovers, E.; Cotecchia, S.; Leurs, R. The effect of mutations in the DRY motif on the constitutive activity and structural instability of the histamine H-2 receptor. *Mol. Pharmacol.* **2000**, *57*, 890−898.
(36) Hamm, H. E. How activated receptors couple to G proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4819−4821.
(37) Halbfinger, E.; Major, D. T.; Ritzmann, M.; Ubl, J.; Reiser, G.; Boyer, J. L.; Harden, K. T.; Fischer, B. Molecular recognition of modified adenine nucleotides by the P2Y(1)-receptor. 1. A synthetic, biochemical, and NMR approach. *J. Med. Chem.* **1999**, *42*, 5325−5337.
(38) Babor, M.; Sobolev, V.; Edelman, M. Conserved positions for ribose recognition: Importance of water bridging interactions among ATP, ADP and FAD-protein complexes. *J. Mol. Biol.* **2002**, *323*, 523−532.
(39) Moro, S.; Jacobson, K. A. Molecular modeling as a tool to investigate molecular recognition in P2Y receptors. *Curr. Pharm. Des.* **2002**, *8*, 2401−2413.
(40) Saenger, W. *Principles of nucleic acid structure*; Springer-Verlag: New York, 1984; Chapter 4, pp 107−110.
(41) Mao, L.; Wang, Y.; Liu, Y.; Hu, X. Multiple intermolecular interaction modes of positively charged residues with adenine in ATP-binding proteins. *J. Am. Chem. Soc.* **2003**, *125*, 14216−14217.
(42) Tsuzuki, S.; Lüthi, H. P. Interaction energies of van der Waals and hydrogen bonded systems calculated using density functional theory: Assessing the PW91 model. *J. Chem. Phys.* **2001**, *114*, 3949−3957.
(43) Reference 40, pp 51−101.
(44) Ravi, R. G.; Kim, H. S.; Servos, J.; Zimmermann, H.; Lee, K.; Maddileti, S.; Boyer, J. L.; Harden, T. K.; Jacobson, K. A. Adenine nucleotide analogues locked in a northern methanocarba conformation: Enhanced stability and potency as P2Y(1) receptor agonists. *J. Med. Chem.* **2002**, *45*, 2090−2100.
(45) Fischer, B.; Boyer, J. L.; Hoyle, C. H. V.; Ziganshin, A. U.; Brizzolara, A. L.; Knight, G. E.; Zimmet, J.; Burnstock, G.; Harden, T. K.; Jacobson, K. A. Identification of potent, selective P2Y-purinoceptor agonists: Structure−activity relationships for 2-thioether derivatives of adenosine 5′-triphosphate. *J. Med. Chem.* **1993**, *36*, 3937−3946.
(46) Filipek, S.; Teller, D. C.; Palczewski, K.; Stenkamp, R. The crystallographic model of rhodopsin and its use in studies of other G protein-coupled receptors. *Annu. Rev. Biomol. Struct.* **2003**, *32*, 375−397.
(47) Hiramoto, T.; Nemoto, W.; Kikuchi, T.; Fujita, N. Construction of hypothetical three-dimensional structure of P2Y(1) receptor based on Fourier transform analysis. *J. Protein Chem.* **2003**, *21*, 537−545.
(48) Koehl, P.; Levitt, M. A brighter future for protein structure prediction. *Nature Struct. Biol.* **1999**, *6*, 108−111.
(49) Flohil, J. A.; Vriend, G.; Berendsen, H. J. C. Completion and refinement of 3-D homology models with restricted molecular dynamics: Application to targets 47, 58, and 111 in the CASP modeling competition and posterior analysis. *Proteins* **2002**, *48*, 593−604.
(50) Vaidehi, N.; Floriano, W. B.; Trabanino, R.; Hall, S. E.; Freddolino, P.; Choi, E. J.; Zamanakos, G.; Goddard, W. A., III. Prediction of structure and function of G protein-coupled receptors. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *99*, 12622−12627.
(51) Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F.; Higgins, D. G. The CLUSTALX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **1997**, *25*, 4876−4882.
(52) Šali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.
(53) Riek, R. P.; Rigoutsos, I.; Novotny, J.; Graham, R. M. Nonalpha-helical elements modulate polytopic membrane protein architecture. *J. Mol. Biol.* **2001**, *306*, 349−362.
(54) Fiser, A.; Kinh Gian Do, R.; Šali, A. Modeling of loops in protein structures. *Protein Sci.* **2000**, *9*, 1753−1773.
(55) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM. A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(56) Woolf, T. B.; Roux, B. Molecular-dynamics simulation of the gramicidin channel in a phospholipid-bilayer. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11631−11635.
(57) (a) Unger, V. M.; Hargrave, P. A.; Baldwin, J. M.; Schertler, G. F. X. Arrangement of rhodopsin transmembrane alpha-helices. *Nature* **1997**, *389*, 203−206. (b) Schertler, G. F. X. Structure of rhodopsin. *Eye* **1998**, *12*, 504−510.
(58) Venable, R. M.; Zhang, Y.; Hardy, B. J.; Pastor, R. W. Molecular dynamics simulations of a lipid bilayer and of hexadecane: An investigation of membrain fluidity. *Science* **1993**, *262*, 223−226.
(59) Nagle, J. F. Area/lipid of bilayers from NMR. *Biophys. J.* **1993**, *64*, 1476−1481.
(60) Steinbach, P. J.; Brooks, B. R. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem.* **1994**, *15*, 667−683.
(61) Connolly, M. L. Solvent-accessible surfaces of proteins and nucleic acids. *Science* **1983**, *221*, 709−713.

(62) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A geometric approach to macromolecular-ligand interactions. *J. Mol. Biol.* **1982**, *161*, 269−288.
(63) Ewing, T. A.; Kuntz, I. D. Critical evaluation of search algorithms for automated molecular docking and database screening. *J. Comput. Chem.* **1997**, *18*, 1175−1189.
(64) Foloppe, N.; MacKerell, A. D., Jr. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **2000**, *21*, 86−104.
(65) Pavelites, J. J.; Gao, J.; Bash, P. A.; Mackerell, A. D., Jr. A molecular mechanics force field for NAD(+), NADH, and the pyrophosphate groups of nucleotides. *J. Comput. Chem.* **1997**, *18*, 221−239.
(66) Møller, C.; Plesset, M. S. Note on an approximate treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618−622.
(67) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(68) (a) Hall, G. G. The molecular orbital theory of chemical valency VIII. A method for calculating ionisation potentials. *Proc. R. Soc, London Ser. A* **1951**, *205*, 541−552. (b) Roothan C. C. J. New developments in molecular orbital theory. *Rev. Mod. Phys.* **1951**, *23*, 69−89.
(69) Ransil, B. J. Studies in molecular structure. IV. Potential curve for the interaction of two helium atoms in single-configuration L.C.A.O.M.O.S.C.F. approximation. *J. Chem. Phys.* **1961**, *34*, 2109−2118.
(70) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.

JM049772M