SUPPLEMENTAL MATERIALS

A STRUCTURAL MODEL OF A LIGAND/GPCR COMPLEX BASED ON EXPERIMENTAL DOUBLE-MUTANT CYCLE DATA: THE MT7 SNAKE TOXIN BOUND TO A DIMERIC hM1 MUSCARINIC RECEPTOR

Catherine Marquer[#3*], Carole Fruchart-Gaillard[#*], Guillaume Letellier[‡], Elodie Marcon[#], Gilles Mourier[#], Sophie Zinn-Justin[‡], André Ménez†, Denis Servent[#2] and Bernard Gilquin[‡1]

From Commissariat à l'Energie Atomique (CEA), Institut de Biologie et Technologies de Saclay (iBiTecS), and CNRS Unité de Recherche Associée 2096, Gif sur Yvette, F-91191, [‡]Service de Bioénergétique, Biologie Structurale et Mécanismes (SB[2]SM), Laboratoire de Biologie Structurale et Radiobiologie, [#]Service d'Ingénierie Moléculaire des Protéines (SIMOPRO), Laboratoire de Toxinologie Moléculaire et Biotechnologie. [*] These authors contribute equally to this work.

Running head: Modeling of the MT7/hM1 dimer muscarinic receptor complex

[1] Corresponding author: Dr Bernard Gilquin, CEA, iBiTecS, Service de Bioénergétique, Biologie Structurale et Mécanismes (SB[2]SM) and CNRS URA 2096, Laboratoire de Biologie Structurale et Radiobiologie, Gif sur Yvette, F-91191, France. Tel: 33 1 69 08 81 53, Fax: 33 1 69 08 47 12, Email: bernard.gilquin@cea.fr

[2] Corresponding author: Dr Denis Servent, CEA, iBiTecS, Service d'Ingénierie Moléculaire des Protéines (SIMOPRO), Laboratoire de Toxinologie Moléculaire et Biotechnologie, Gif sur Yvette, F-91191, France. Tel: 33 1 69 08 52 02, Fax: 33 1 69 08 90 71, Email: denis.servent@cea.fr

[3] Present address: Centre de Recherche de l'Institut du Cerveau et de la Moelle, UPMC, INSERM UMR-S975, Centre National de la Recherche Scientifique (CNRS) UMR7225, Hôpital de la Pitié-Salpêtrière, Paris, France.

Supplemental Table 1: Characteristics of the five docking runs performed to model the MT7/hM1 dimer complex on the basis of experimental double-mutant cycle data

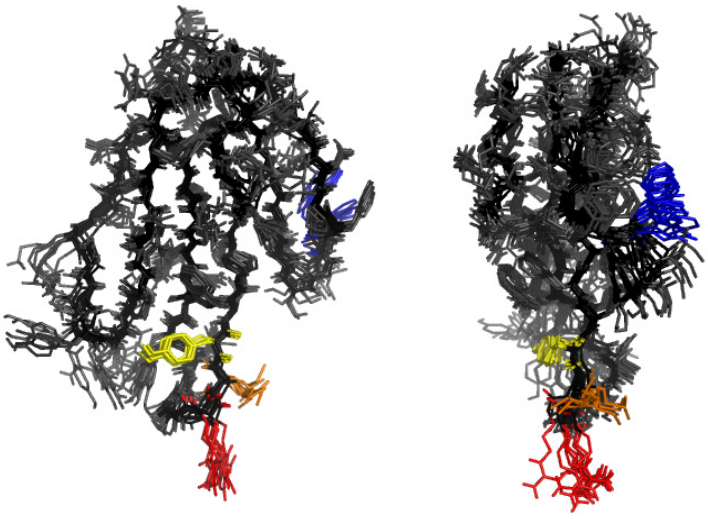
	Docking run name	Nb restraints[1]	Nb no-contact[2]	Method for setting the initial orientations of hM1 protomers	D-init[3] (Å)	Sampling of the initial orientations of the two hM1 protomers	Rotation of the hM1 around the Z axis (range; increment)	Namb hM1[4]	Number of initial generated configurations	Number of docked structures	Number of docked symmetric structures (Sdev < 2.0)
1	"random"	33	28	CHARMM coor orient	45	random rotation	random rotation	3	9720	8014	1272
2	"symmetric"	33	28	CHARMM coor orient	35	symmetric rotation	systematic rotation 0-360 ; 10	3	11664	9780	5933
3	"TM6/TM7_sym"	33	28	CHARMM coor orient	30	symmetric rotation	systematic rotation 80-120 ; 2	16	11664	10755	8700
4	"TM6/TM7_without"	33	28	CHARMM coor orient	30	symmetric rotation	systematic rotation 80-120 ; 2	0	11664	9287	7739
5	"TM6/TM7_cxcr4"	33	28	align on cxcr4 with pymol command align	35	symmetric rotation	systematic rotation 65-135 ; 2	0	11664	9248	6313
	Total number of complexes								**56376**	**47084**	**29957**

1 Nb restraints: number of ambiguous restraints derived from the double mutant data (ΔΔGint > 0.7 kcal/mol) and used during the docking procedure

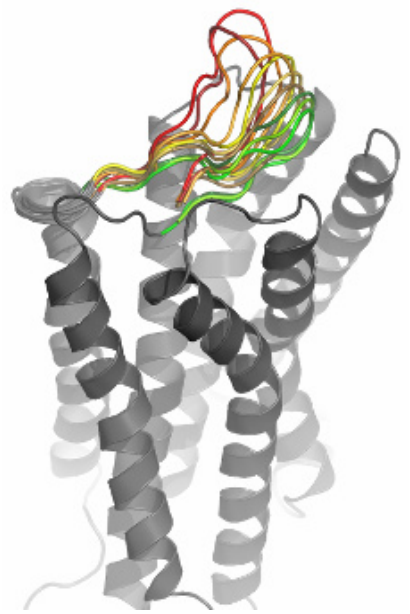
2 Nb no-contact: number of uncoupled residue pairs identified from double mutant data (ΔΔGint < 0.7 kcal/mol) and used to validate the models

3 D init: initial distance between the centers of mass of the two hM1 protomers (Å).

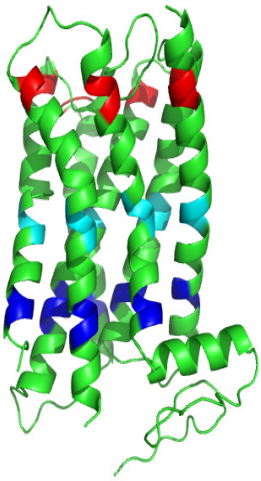
4 Namb hM1: number of ambiguous restraints between the two hM1 protomers.



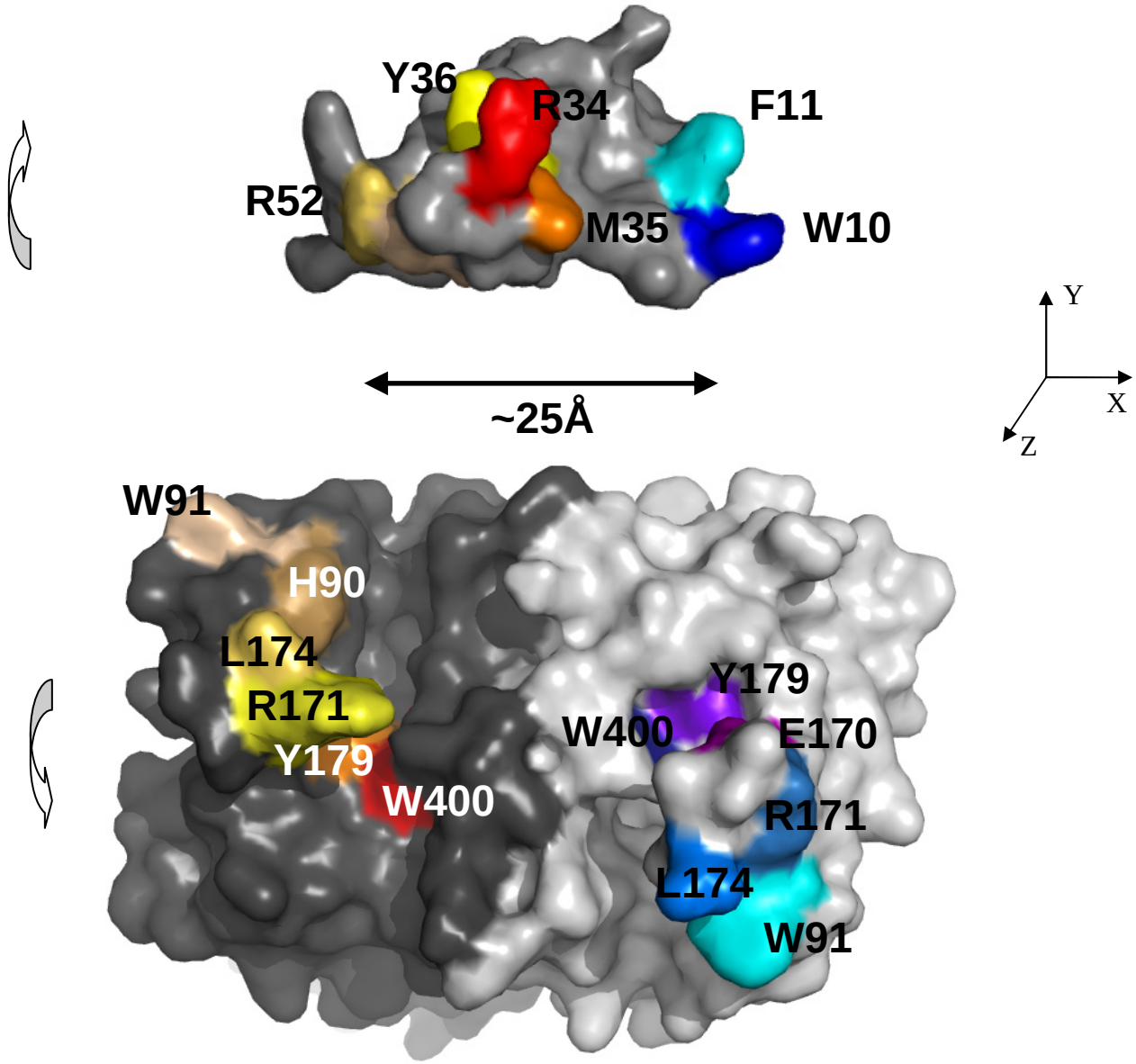
Supplemental Figure S1: Two orthogonal views of the 9 superimposed frames extracted from the MT7 molecular dynamics simulation and used in the docking. Hot spot residues R34, M35, Y36 and W10 are colored in red, orange, yellow and blue, respectively.



Supplemental Figure S2: Ten different conformations of the E2 loop generated by the activated molecular dynamics simulation. Transmembrane regions are superimposed. The E2 loop is colored from green, the initial conformation at 0 ns, to red, the conformation at 1 ns.



Supplemental Figure S3: **Residues involved in ambiguous distance restraints between the two subunits of the hM1 dimer.** Three ambiguous restraints were introduced. The Ca atoms involved in the first, second and third restraints are colored in red, cyan and blue, respectively. For details, see Supplemental Experimental Procedures.



Supplemental Figure S4: Visualization of the double mutant data on the MT7 / hM1 dimer interface. MT7 and hM1 dimer are displayed on the top and bottom views, respectively. In order to visualize the interface, MT7 and hM1 dimer are separated by a translation along the Z-axis and rotated by 90° around a X-axis passing though their center of mass. Therefore, MT7 is displayed with R34 (in red) pointing towards the observer and the hM1 dimer presents its extracellular surface. Mutated surface residues were colored in a range of different colors from red to light yellow ($hM1_A$/MT7 interface) and dark blue to light blue ($hM1_B$/MT7 interface). The change in color from dark to light reflects the increase of the intensity of coupling between this residue and a partner residue. R52, which belongs to the $hM1_A$/MT7 interface and which is strongly coupled to Y179, is consistently represented in yellow-orange.

Supplemental Table 2: Structural statistics

Whole complex	
Total number of residues	579
Deviation from standard geometry ([1])	
RMS-deviation in bond distances (Å)	0.011 ± 2e-10
RMS-deviation in bond angles (°)	2.07 ± 0.02
Standard deviation of omega values (°)	6.47 ± 0.4
Van der Waals energy (kcal/mol)	-2651.23 ± 16.9
Electrostatic energy (kcal/mol)	-8296.28 ± 43.6
Ramachandran plot statistics ([1])	
Residues in most favoured region	86.7 %
Residues in additionally allowed region	11.8 %
Residues in generously allowed region	0.9 %
Residues in disallowed region	0.5 %
Ligand receptor interface MT7/hM1 dimer	
Van der Waals energy (kcal/mol)	-85.91 ± 6.7
Electrostatic energy (kcal/mol) ([3])	-128.84 ± 30.4
Econst (kcal/mol)	25.96 ± 2.1
Buried accessible surface area (Å^2) ([4])	2390 ± 220
Dimer interface	
Van der Waals energy (kcal/mol)	-87.66 ± 9.0
Electrostatic energy (kcal/mol) ([3])	-98.94 ± 35.5
Buried accessible surface area (Å^2) ([4])	3780 ± 310

([1]) Calculated with WHAT-IF
([2]) Calculated with PROCHECK (1)
([3]) Electrostatic energy was calculated with CHARMM using parameter charmm19 and distant dependant dielectric constant (rdie) (2).
([4]) The accessible surface area was calculated with CHARMM with a probe radius of 1.4 Å. The buried surface area was calculated as the difference between the accessible surface area (ASA) for the uncomplexed partner and the same accessible surface area in the complex (for ligand-receptor interface: ASAcomplex – ASAMT7 – ASAdimer ; for dimer interface: ASAdimer – $ASAhM1_A$ – $ASAhM1_B$).

Supplemental Experimental Procedures

cDNAs encoding mutant hM_1 and hM3 receptors

The human M_1 and M_3 receptors were expressed from the m1 and m3 genes (kindly provided by Prof. Buckley; University of Leeds, UK). To construct the chimeric receptors, the m1 and m3 genes were subcloned into the mammalian expression vector pcDNA3.1(-) (Invitrogen). The mutations for punctual mutations or for the construction of chimeric receptors 1, 3, 4 and 6 were introduced into the M_1 or M_3 receptors by site-directed mutagenesis using the QuikChange™ kit (Stratagene). The constructs coding for chimeric receptors 2 and 5 were created by substituting a specific segment of E2 of m1pcDNA3.1(-) with the homologous segment of E2 of m3pcDNA3.1(-) and *vice versa*. Restriction sites were introduced into M_1 and M_3 receptors at the following conserved amino acid residues by site-directed mutagenesis: TM4, *BspeI* site introduced at alanine-160 (M_1) and alanine-203 (M_3); TM5, *ECoRV* site introduced at threonine-192 and alanine-193 (M_1) and alanine-237 (M_3). Constructs were sequenced to ensure that no other change was introduced in the amino acid sequence. Chimeric receptors 2 and 5 were constructed by subcloning. In order to remove the mutations inserted with the restriction sites *BspeI* and *ECoRV*, an additional mutagenesis step was performed by Multi site-directed mutagenesis QuikChange™. The presence of the desired mutations was confirmed by DNA sequencing using an ABI PRISM™ A310 Genetic Analyses (PerkinElmer) sequencer. The plasmids obtained were amplified in *Escherichia coli* XL1Blue using the midi protocol of the Wizard® plasmid purification kit from Promega.

Stable and transient mAChRs expression

Heterologous transient expressions of wild-type and punctually mutated hM_1 receptors and chimeras were performed in COS cells cultured in 10-cm-diameter tissue culture dishes (Falcon, Cowley, UK). COS cells were incubated at 37°C in an atmosphere of 5% CO_2 and 95% humidified air in Dulbecco's Modified Eagle Medium supplemented with 10% fetal calf serum, 1% L-glutamine and 1% penicillin/streptomycin. All products were from Sigma. Cells were transfected with 10 μg of cDNA encoding hM_1 mutants with the calcium phosphate methodology, as described previously (3). Three days after transfection, cells were harvested using Versen buffer (PBS with 5mM EDTA), washed with ice-cold phosphate buffer and centrifuged at 1700 g for 3 min (4°C) three times. The pellet was stored at -20°C. For membrane preparation, the cell pellet was suspended in ice-cold buffer (1 mM EDTA, 25 mM Na phosphate, 5 mM $MgCl_2$, pH 7.4) and homogenized using an Elvehjem-Potter homogenizer (Fisher Scientific Labosi). The homogenate was centrifuged at 1700 g for 15 min at 4°C. The sediment was re-suspended in buffer, homogenized and centrifuged at 1700 g for 15 min at 4°C. The combined supernatants were centrifuged at 35000g for 30 min at 4°C and the pellet was suspended in the same buffer (0.1 ml/dish). Protein concentrations were determined according to the Lowry method using bovine serum albumin as standard. Membrane preparations were aliquoted and stored at -80°C.

Molecular Dynamics Simulations of MT7

Molecular dynamics simulation of MT7 in water was performed using periodic boundary conditions. The molecule was immersed into a box of pre-equilibrated TIP3 water molecules (4) with at least 10 Å of water between the protein and the edges of the box. Water molecules overlapping the protein, more precisely those having their oxygen atoms within a distance of 2.8 Å from any heavy atom of the protein, were removed. The size of the box was set to 65x60x50 Å resulting in a system containing 6567 water molecules. The system was minimized in the presence of harmonic restraints on the protein heavy atoms in order to preserve the global conformation while optimizing water–protein interactions. After heating and equilibration of the system, 2.5 ns of simulation at 300K were performed. 1000 frames were extracted from the trajectory and clustered into 9 groups. Finally, one frame of each cluster was selected (Supplemental Figure S1). The calculations were performed with CHARMM (2) package version c32b1 and force fields CHARMM22 (5) on a local cluster of 22 AMD

Opteron processors. All bond lengths of hydrogen atoms were constrained using the SHAKE (6) algorithm using a time step of 1 fs.

Structural model of hM1 receptor

Sequences of the hM1 receptor and bovine rhodopsin were aligned with respect to the "pin-points" as described by Baldwin et al (7). The percentage of homology deduced from this alignment is 13% on the whole structure and rises up to 24% for the helix core. A model of the muscarinic receptor was generated with MODELLER version 8.2 (8) using rhodopsin X-ray structure solved at 2.2 Å (pdb code 1U19) (9) as template. The long hM1 intracellular loop 3 not present in the rhodopsin template was removed. As the sequences of extracellular loop 2 (residues 169-177) and loop 3 (residues 391:396) show no similarity with the template, they had to be further refined. The tip (residues 167-174) of extracellular loop 2 was predicted *ab-initio* with RAPPER (10). For extracellular loop 3, a good template was found in the PDB (pdb code 1KQF). PROCHECK was used to evaluate the quality of the model and shows less than 1% of residues in disallowed Ramachandran plot regions (1).

Procedure used in the docking calculations of MT7 on hM1 dimer

Starting conformations of MT7 and hM1- Nine different conformations of MT7 and 36 different conformations of hM1 in which the E2 loop adopts different structures were combined in order to create initial unbound complexes. For each unbound complex, docking was repeated 30 times for the "random_fix" docking run and 36 times for all the others docking runs. Thus, 9720 starting complexes were generated for the "random fix" run and 11664 for all the other docking runs. (Table S1, col 9).

Location of MT7 - MT7 was oriented with the axis defined by Arg34.Cα-Arg34.Nζ directed towards the hM1 receptors. The center of mass of MT7 was located between the two hM1 receptors, at 7Å far from the extracellular face of the receptors. MT7 was randomly rotated along the Arg34.Cα Arg34.Nζ axis to generate different orientations relatively to the hM1 receptors.

Four dimension restraint molecular dynamics - Four dimension molecular dynamics under distance restraints was performed at 500K. Atomic displacements along a fictional fourth dimension were allowed to reduce the core-core repulsion between toxin and receptor, thus enabling easier structural rearrangements of the side chains at the protein-protein interface (11). The position in the fourth dimension was set to 1, 0, -1 Å for respectively MT7, the first hM1 receptor and the second one respectively. During the first 4000 steps the force constant on the ambiguous restraints gradually increased. Next, 20000 steps of production run were performed to sample the relative positions of the MT7 on hM1 and optimize the interface between the side-chains of both partners.

Loop E2 - In the first steps of the docking, to maximize the flexibility of the E2 loop of both receptors, the two peptide bonds between Gly169-Glu170 and Gln177-Cys178 that links E2 to TM4 and TM5 were replaced by a distance restraint with a strong constant force equal to 200 kcal/mol/Å^2. Then, the peptide bonds between Gly169-Glu170 and Gln177-Cys178 that links E2 to TM4 and TM5 were re-introduced. After re-introducing the Gly169-Glu170 and Gln177-Cys178 bonds, 10000 additional steps were performed. During the last 30000 steps, the temperature decreased from 500K to 300K, and the protein was brought back to regular three-dimensional space (The fourth dimension position decreased from its initial value to 0 Å). Lastly, the structure was minimized during 250 steps.

Harmonic restraints - To preserve the structure of both hM1 receptors during the docking, the backbone atoms C, N and Cα and the Cβ atoms (except those of loop E2: residues 169-177) were kept fixed for the first receptor and for the second one maintained by the CHARMM BESTFIT command. The MT7 backbone was treated as a rigid body by applying restraints on atoms C, N, Cα and Cβ with the CHARMM BESTFIT command. The CHARMM BESTFIT command imposes harmonic restraints (100 kcal/mol/Å^2) on the selected atoms by allowing complete freedom of translation and rotation of MT7. Therefore only the extracellular loops of each receptor and the side-chains of both partners were free.

Ambiguous distance restraints between the two subunits of hM1 dimer - In the two docking run procedures ("random" and "symmetric" Table 1), we introduced three ambiguous restraints favoring the proximity between the upper, the central and the lower parts of the two receptors. These three ambiguous distance restraints involved several residues of each TM. Therefore these did not favor any interface between the two protomers. These restraints were design in order to favor dimer conformation with parallel subunits. Several CA atoms were chosen on each TM. The ambiguous distance restraints between protomer upper parts involved CA atoms of residues 27-29(TM1), 84-86(TM2), 96-98(TM3), 164-166(TM4), 185:186(TM5), 387:389(TM6), 400-402(TM7) (in red); between protomer middle parts 38-40(TM1), 70-73(TM2), 110-111(TM3), 152-153(TM4), 199-201(TM5), 373-375(TM6), 410-412(TM7) (in cyan); between protomer lower parts 46-48(TM1), 63-65(TM2), 118-121(TM3), 143-145(TM4), 207-209(TM5), 365-367(TM6), 418-420(TM7) (in blue) (supplemental Figure S3).
For the run "TM6/TM7 sym", we used a different method to define the ambiguous distances between the two protomers and we increased the number of these restraints. 16 ambiguous restraints were introduced to improve the hM1-hM1 interface. The selection of the CA atoms involved in the ambiguous distance restraints was based on the Z-position of CA atoms of each monomer: 16 Z values regularly spaced between 16 and -16 were chosen. For each Z value CA atoms located at this Z level +/- 1 Å were selected on each protomer and an ambiguous restraint between them was introduced.

Ambiguous distance restraints used to drive the docking - The restraints were introduced as NMR restraints via the CHARMM NOE command with the MINDIST option to ensure that the restraint would be active only between the nearest pair of atoms of the residue pair. The restraint potential was a square well function in which the value of RMIN was set to 1.8 Å. The ambiguous restraints were split in three groups: i) First group of restraints involved hot spot residues (MT7 Arg34 and hM1 Trp400 & Tyr179). For this group RMAX was set to 2 Å and the constraint force KMAX increased from 3 to 12 kcal/mol/$Å^2$. ii) Second group: restraints involved MT7 and hM1 residues except those belonging to E2. For this group RMAX was set to 4 Å and the constraint force KMAX increased from 1 to 4 kcal/mol/$Å^2$. iii) Third group of restraints between MT7 residues and hM1 residues belonging to E2 loop. For this group RMAX was set to 6 Å and the constraint force KMAX increased from 0 to 1 kcal/mol/$Å^2$.

REFERENCES

1. Laskowski, R. A., M W MacArthur, D S Moss and J M Thornton. (1993) *J. Appl. Cryst.* **26**, 283-291
2. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem.* **4**, 187-217
3. Chen, C., and Okayama, H. (1987) *Mol Cell Biol* **7**, 2745-2752
4. Jorgensen, W. L. (1981) *J. Am. Chem. Soc* **103**, 335-340
5. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evansek, J. D., Field, M. J., Fischer, S., Gao, J., Guo, S., Ha, S., and al, e. (1998) *J. Phys. Chem.* **B102**, 3586-3616
6. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. (1977) *J. Comp. Chem.* **23**, 327-341
7. Baldwin, J. M., Schertler, G. F., and Unger, V. M. (1997) *J Mol Biol* **272**, 144-164
8. Sali, A., and Blundell, T. L. (1993) *J Mol Biol* **234**, 779-815
9. Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P., and Buss, V. (2004) *J Mol Biol* **342**, 571-583
10. Depristo, M. A., de Bakker, P. I., Johnson, R. J., and Blundell, T. L. (2005) *Structure* **13**, 1311-1319
11. van Schaik, R. C., Berendsen, H. J., E., T. A., and F., v. G. W. (1993) *J Mol Biol.* **5**, 751-762