Oxidation state dependent conformational changes of HMGB1 regulates the formation of the CXCL12/HMGB1 heterocomplex

Enrico M. A. Fassi[1•], Jacopo Sgrignani[1•], Gianluca D'Agostino[1], Valentina Cecchinato[1], Maura Garofalo[1,3], Giovanni Grazioso[4], Mariagrazia Uguccioni[1,5*] and Andrea Cavalli[1,2*]

[1] Institute for Research in Biomedicine, Università della Svizzera Italiana, CH-6500 Bellinzona, Switzerland

[2] Swiss Institute of Bioinformatics, Lausanne, Switzerland

[3] University of Lausanne (UNIL), CH-1015, Lausanne, Switzerland.

[4] Dipartimento di Scienze Farmaceutiche, Università degli Studi di Milano, Milan, Italy

[5] Humanitas University, Department of Biomedical Sciences, 20090 Pieve Emanuele – Milan, Italy

[•] these authors equally contributed to the work.

* these authors equally contributed to the work.

Corresponding authors

Dr. Andrea Cavalli

Institute for Research in Biomedicine

Università della Svizzera Italiana

Via Vincenzo Vela 6

CH-6500 Bellinzona, Switzerland

Email: andrea.cavalli@irb.usi.ch

Prof. Mariagrazia Uguccioni

Institute for Research in Biomedicine

Università della Svizzera Italiana

Via Vincenzo Vela 6

CH-6500 Bellinzona, Switzerland

Email: mariagrazia.uguccioni@irb.usi.ch

Abstract

High-mobility Group Box 1 (HMGB1) is an abundant protein present in all mammalian cells and involved in several processes. During inflammation or tissue damage, HMGB1 is released in the extracellular space and, depending on its redox state, can form a heterocomplex with CXCL12. The heterocomplex acts exclusively on the chemokine receptor CXCR4 enhancing leukocyte recruitment.

Here, we used multi-microsecond molecular dynamics (MD) simulations to elucidate the effect of the disulfide bond on the structure and dynamics of HMGB1.

The results of the MD simulations show that the presence or lack of the disulfide bond between Cys23 and Cys45 modulates the conformational space explored by HMGB1, making the reduced protein more suitable to form a complex with CXCL12.

Introduction

High-mobility Group Box 1 (HMGB1) is an abundant chromatin-associated protein present in all mammalian cells. It is formed by 215 amino acids, divided into two domains, "BoxA" (Gly2-Ile79) and "BoxB" (Phe89-Arg163), connected by a nine amino acid loop, and a highly disordered negatively charged C-terminal tail. BoxA contains a pair of cysteines (Cys23 and Cys45) that can form a disulfide bond under oxidative conditions. In contrast, only one unpaired cysteine is present in BoxB (Cys106, Figure 1A).[1-2]

Depending on its cellular localization, HMGB1 performs different functions. In fact, as a nuclear protein, it is involved in DNA repair, transcription, telomere maintenance, and genome stability,[2,3, 4] while during cellular death or inflammation, HMGB1 is released in the extracellular space where it functions as an alarmin.[5-6]

According to multiple studies, several HMGB1 functions depend on its redox states.[7-8] The nuclear and cytosolic environments are characterized by a negative redox potential that maintains HMGB1 in reduced form (fr-HMGB1). During an inflammatory process, the extracellular space, enriched in reactive oxygen species, might lead to the formation of a disulfide bond between cysteines at positions 23 and 45 of BoxA (ds-HMGB1).[9] ds-HMGB1 activates Toll-like Receptor 2 (TLR2) and 4 (TLR4) inducing the release of proinflammatory chemokines and cytokines activating innate and adaptive immune responses. On the contrary, fr-HMGB1 binds to the receptor for advanced glycation end products (RAGE), modulating autophagy.[6, 10, 11]

The CXC ligand 12 (CXCL12) is expressed in many tissues both under homeostatic and inflammatory conditions and can stimulate cellular recruitment by activating the CXC chemokine receptor type 4 (CXCR4).[12] In 2012, researchers in our group have shown the enhancement of the activities induced by the interaction of CXCL12 with

its specific receptor CXCR4 in the presence of HMGB1.[13] In particular, a suboptimal concentration of CXCL12, which per se would not trigger any chemotactic response, efficiently promotes migration of human monocytes, by forming a heterocomplex with fr-HMGB1.[13, 14] More recently, other studies demonstrated the important role of the heterocomplex in tissue regeneration[10, 15, 16] and in fueling the inflammatory response in patients with Rheumatoid Arthritis.[17]

A particular feature of the CXCL12/HMGB1 heterocomplex is that only fr-HMGB1 can complex with CXCL12, promoting CXCR4-induced response.[14] This appears contradictory because the extracellular space, where the heterocomplex is formed, is rich in reactive oxidative species[14]. However, under specific conditions, cells can release glutathione reductase and enzymes of the thioredoxin system to counteract the oxidative stress in the microenvironment, contributing to maintain HMGB1 in the reduced state.[17, 18]

While a structure of the heterocomplex is currently unavailable, NMR chemical shift mapping clearly showed an interaction between CXCL12 and the two domains of HMGB1 (BoxA and BoxB), separately.[13] Furthermore, the same experiments showed that the binding of CXCL12 to HMGB1 induces conformational changes in the *N*-terminal domain of CXCL12 which is required to trigger the activation of the receptor. Based on these data, it was hypothesized that the heterocomplex is formed by two CXCL12 molecules bound to fr-HMGB1 (one to BoxA and one to BoxB), and that the heterocomplex would bind CXCR4 dimers (Figure 1B).[13]

In this study, aiming to validate the supposed mode of action of the heterocomplex, we applied several molecular modeling techniques. such as molecular dynamics (MD) simulations and protein-protein docking, to investigate which structural and/or conformational differences between the two redox states of HMGB1 could explain the

different affinity of fr- and ds-HMGB1 for CXCL12.

According to our findings, ds-HMGB1 tends to be more compact and display a lower accessible surface than fr-HMGB1, while the structure of BoxA remains essentially unchanged in the two states. Furthermore, in-depth analysis of the simulations and the results of protein-protein docking calculations showed that the vast majority of the conformations assumed by fr-HMGB1 are able to bind two CXCL12 molecules with an orientation and distance optimal to trigger the activation of CXCR4 dimers. We, therefore, propose that functional differences between fr- and ds-HMGB1 are at least partially caused by global changes in the configurational landscape of HMGB1.

Methods

Systems setup and MD simulations.

MD simulations are a powerful tool already applied to the study of some mechanistic aspects of the HMGB1 cellular functions.[19, 20]

In this case, the HMGB1 structure solved by NMR spectroscopy (PDB ID 2YRQ), was used as a starting point for the simulations. As the first residue (Met1) of the protein is cleaved during posttranscriptional processing,[21] this amino acid was deleted from the model and only the region from Gly2 to Arg170 (i.e., BoxA, BoxB, and the connecting loop) was considered in the MD simulations. All the investigated HMGB1 models (fr- of ds-) were first minimized using the program ALMOST.[22] Then, the TLEAP module of AmberTools16 was used to solvate the protein in a box of water with a minimum distance of 10 Å from the protein surface. The net charge of the system was neutralized by adding a proper number of ions (17 Cl^- ions). The ff14SB[23] force field parameters were used to describe the protein, while the TIP3P[24] model and the parameters proposed by Joung et al.[25] were used for water and counter ions, respectively. The solvated system

was relaxed by a two-step protocol to remove atomic clashes.[26] First, we performed an energy minimization for 10,000 steps, or until the energy gradient of 0.2 kcal/mol/Å was reached, restraining the atomic coordinates of backbone with harmonic potential (k=20 kcal/mol/Å^2). This first phase was followed by an energy minimization for 100,000 steps or until an energy gradient of 0.0001 kcal/mol/Å was reached, without any restraint. After minimization, the temperature of the system was gradually increased to 300 K over 40 ps under constant volume condition (NVT) constraining the backbone coordinates in the first 20 with a harmonic potential (k=20 kcal/mol/Å^2). Finally, the system was equilibrated at 300 K for 20 ps under constant pressure conditions (NPT, 1 atm). Pressure and temperature were maintained constant using the Berendsen barostat and thermostat, respectively.[27] Electrostatic interactions were treated with PME[28] with a cutoff of 9 Å. During the calculations, all bonds involving hydrogen atoms were constrained with the SHAKE[29] algorithm. All calculations were performed using the PMEMD of Amber16 code in the GPU accelerated version[30] with a time step of 2 fs.

Production runs were carried out using the following scheme. After the first simulation of 1 μs, 29 of the saved frames were randomly selected and used as a starting point for 29 additional simulations (see Table 1). The atom velocities were reassigned at the beginning of each simulation to obtain uncorrelated and independent trajectories.

Trajectory Analysis

The protein radius of gyration (RoG) was computed using the cpptraj[31] module available in AmberTools16 including all the protein residues. To assess the convergence of RoG calculation, 75000 snapshots sampled over the 30x1 μs trajectories were divided into six groups of 12500 snapshot. Then the snapshots belonging to one of the six groups

were excluded from the calculation and the results compared with those obtained using the full conformation ensemble (Figure S2).

The RMSDs of BoxA (Lys8 to Ile79) and BoxB (Lys96 to Arg163) were computed with the VMD[32] software, using the first conformation from the HMGB1 NMR bundle (PDB ID code 2YRQ) as a reference.

The solvent accessible surface area (SASA) was computed for the entire protein, BoxA, and BoxB using the LCPO algorithm[33] implemented in the cpptraj module of Amber16. Finally, atom-atom and residue-residue contact analyses were carried out using the g_contacts program developed by Bau and Grubmuller.[34] Given that 1H-1H NOEs are detectable up to a distance of approximatively 5-6 Å, we used a cut-off of 6 Å in the contacts analysis.

Clustering procedure

The sampled protein conformations were clustered with the g_cluster (GROMOS method) program available in the GROMACS software package (version 5.1.2).[35] After several clustering runs (Table S2) and an accurate visual inspection of the results, we verified that the application of an RMSD cutoff of 1.4 nm permitted us to discriminate different system conformations and to limit the number of singleton clusters simultaneously.

Twelve and eleven clusters were obtained for fr- and ds-HMGB1, respectively. For both systems, the centers of the first three clusters, which in both cases accounted for more than 90% of the sampled conformations, were selected for further analysis.

Docking procedure

The centers of the three most populated clusters fr- and ds-HMGB1 were then used in

docking calculations to obtain the putative structures of the $CXCL12_2$/HMGB1 heterocomplex.

For CXCL12, we used the center of the most populated cluster (75.2 % of the sampled structures) obtained by clustering (RMSD cutoff 3.5 Å) the simulation of 300 ns, carried out starting from the NMR structure deposited in the PDB databank with the PDBID 2KEC.[36] MD simulations were performed with the same setup and force field parameters previously used for HMGB1, adding disulfide bonds between the pairs of cysteine residues at positions 9-34 and 11-50.

Docking calculations were performed using the HADDOCK 2.2 webserver.[37] These calculations require the user to define the residues forming the binding site and, while the residues involved in the interaction between the BoxB and CXCL12 have been identified by NMR chemical shift perturbations and reported in our previous study,[13] the residues forming the BoxA binding site have not yet been defined. Therefore, for BoxA we used 'homologous' residues obtained aligning the structures of both HMGB1 boxes (Table 2). Only the structures of the complex with the best HADDOCK scores were kept for further analysis.

MD simulations of the $CXCL12_2$/HMGB1 *complexes*

The structures of the heterocomplex obtained by docking calculations were prepared and simulated for 500 ns with the same parameters and set-up used for HMGB1 and CXCL12.

During the first 200 ns, a harmonic distance restraint was applied between the centers of mass of HMGB1 and CXCL12 to optimize atomic contacts at the protein-protein interface. In particular, the force constant (k) was slowly decreased from 400 kcal/mol/$Å^2$ to 0 over the first 200 ns. Then the systems were simulated for additional

300 ns. To increase the statistical significance of the calculations these simulations were repeated three times.[38]

Analysis of the trajectories of the $CXCL12_2$/HMGB1 *complexes*

The last 300 ns of the MD simulations trajectories computed for the $CXCL12_2$/HMGB1 complexes were first visually analyzed to assess the stability of the complex.

Then the distance between the *N*-terminal domains of the two CXCL12 molecules were computed with the aim of determining whether the obtained $CXCL12_2$/HMGB1 complexes conformations could potentially bind to and activate CXCR4 dimers.

The distance between the two binding sites in the CXCR4 receptor dimers served as the reference value. This value was determined measuring the distance between the two chemokine *N*-terminal domains (Cα of Leu1) in the structure of a CXCR4 receptor (pdb code 4RWS[39]) in complex with a CXCL12 analog (viral macrophage inflammatory protein II (vMIP-II)).

The dimer structure was obtained applying the crystal symmetry to the deposited structure (Figure S4).

Results and discussion

HMGB1 MD simulations

According to experimental observations, only fr-HMGB1 can form a heterocomplex with CXCL12, enhancing its chemotactic activity.[13, 15]

These experimental findings can be explained by different hypotheses. Indeed, the making/breaking of the disulfide bond can: (1) influence the local structure of BoxA making it unable to bind CXCL12, (2) induce a shift of the protein conformational ensemble making the HMGB1 less suitable to form the heterocomplex or, (3) the observed effect is due to a combination of the above factors.

Aiming to provide insights into the factors that regulate the formation of the heterocomplexes, we simulated both the systems (fr- and ds-HMGB1) for 30 μs MD and analyzed the outputs, focusing on proper descriptors such as the solvent accessible surface area (SASA, Figure 2A), radius of gyration (RoG, Figure 2B) and RMSD with respect to NMR structure (PDB ID code 2YRQ, Figure S1).

NMR studies on ds-HMGB1, performed by Wang et al.[40] highlighted a set of 1H-1H NOE signals due to the interaction of Phe38 with Val20, Gln21, and Arg24 not detected for fr-HMGB1. As a consequence, a different orientation is assumed by Phe38 in the available HMGB1 structure (pdb codes: 2YRQ and 2RTU, Figure 2).

We subsequently focused our attention on descriptors (distances, residue-residue, and atom-atom contacts) capable to capture the differences in the structure and dynamics of this region in the two different oxidation states (Figure 2 and Table S1).

RMSD analysis of BoxA (Figure S1C) resulted in very similar values for both ds-HMGB1 and fr-HMGB1, indicating that the formation of the Cys23-Cys45 disulfide bond in BoxA does not strongly alter the local conformation.

Concerning the Phe38 orientation, considering that 1H-1H NOE signals origin by short range interactions (< 5-6 Å), we monitored both the distribution of the distances between the center of mass Phe38 and the three interacting residues indicated by the NMR experiments (Val20, Gln21 and Arg24) and the percentage of the simulation time in which the atom-atom contacts responsible for the 1H-1H NOE signals are present (Table S1).

This analysis (Figure 2 E, F, G and Table S1) confirmed that the presence of the disulfide bond facilitates the interaction of Phe38 with Val20, Gln21 and Arg24 however, the results of both residue-residue distance analyses and atom-atom contacts suggest that, in agreement with the dynamical nature of the system, Phe38 can flip

between different conformation in both fr- and ds-HMGB1.

The RoG analysis (Figure 2A) showed a difference between the conformational spaces visited by the two systems. While two separate peaks are visible for fr-HMGB1 (the first centered at ~24 Å and the second at ~34 Å), only the first peak is clearly visible for ds-HMGB1. Based on this observation, the system containing the disulfide bond more frequently assumes a more compact conformation than fr-HMGB1.

Finally, the SASAs for the entire protein (Figure 2B) and for BoxA and BoxB (Figure S1A-B), were estimated to evaluate the propensity of the two different HMGB1 forms to bind CXCL12. In all cases, we obtained a lower value for ds-HMGB1 than fr-HMGB1.

Summarizing, all the analyses of the simulations indicate that the presence or absence of the disulfide bond modulates the protein size, the reciprocal orientation of both the boxes and the SASA of HMGB1 without significantly altering the structure of BoxA and BoxB. As a consequence, a change in the conformational space explored by ds- or fr-HMGB1 seems to be the molecular determinant of the reduced fr-HMGB1 propensity to form a complex with two CXCL12 molecules reported in experimental studies.[13-15]

$CXCL12_2$-HMGB1 binding

To further investigate the propensity of the two HMGB1 redox states to bind CXCL12, protein-protein docking studies were performed. Representative structures were selected from the protein ensembles obtained by MD simulations by cluster analysis.

In the case of fr-HMGB1, the two most populated clusters (Figure 3A and 3C) include 55% and 20% of the conformations sampled by the system during MD simulations. Importantly, in both cluster center structures, the two CXCL12 binding sites are free

(i.e., not interacting with other protein regions) and potentially able to bind CXCL12, with the N-terminal domain oriented in the same direction.

In contrast, the representative conformation (cluster center, Figure 3E) from the third cluster, which comprises the 14% of the generated conformational ensemble, is more compact, with the two domains interacting and, consequently, unable to bind CXCL12.

For ds-HMGB1, we observed an almost reversed trend. In this case, the first and the third most populated clusters (Figure 3G and 3K) contain 54% and 13% of the conformations, respectively. Interestingly, in both cluster centers, BoxA and BoxB are involved in reciprocal interactions that significantly limit or nullify their abilities to bind one or more CXCL12 molecules.

Only the representative conformation (center cluster) from the second cluster (Figure 3I), which accounts for 25% of the total conformations, is expanded and both domains are available to bind one CXCL12 molecule.

In summary, considering the entire conformational ensemble of fr- and ds-HMGB1 sampled during 30 μs of MD simulations, we can estimate that while the ~75% of the conformations assumed by fr-HMGB1 can activate the CXCR4 dimers, only ~25% of the observed ds-HMGB1 conformation can do the same.

Docking calculations were performed to investigate which of the cluster centers were able to bind two CXCL12 molecules and obtain putative structures of the CXCL12/HMGB1 heterocomplexes (Figure 3B, D, F, H, J, and L). These calculations confirmed our findings from the analysis of the MD simulations trajectories. In particular, CXCL12 could be docked in the correct binding site only in the two center structures from the first two clusters from the simulations of fr-HMGB1 (fr-HMGB1(I) and fr-HMGB1(II)). Moreover, in this case, the two *N*-terminals domains are oriented in the same direction, and then the formed heterocomplexes have an optimal

conformation to bind a CXCR4 dimer. In contrast, the third cluster center structure fr-HMGB1 (III) is unable to bind two CXCL12 molecules due to the inaccessibility of BoxA.

In the case of ds-HMGB1, the docking two chemokines in the correct binding site, was only possible with the structure of the second cluster center. However, a visual inspection of the resulting complex (Figure 3L) reveals that the N-terminal domains of the two CXCL12 are not oriented in the same direction, making impossible the activation of CXCR4 dimers.

Docking calculations were performed using static structures, thus completely neglecting protein dynamics and the reciprocal induced fit effects. Therefore, aiming to explore the stability of the complexes obtained by docking, we simulated them for 500 ns. It should be noted that these simulations were not performed to fully explore the conformational ensemble of the complex, but to relax the system and obtain more reliable models

The simulations were analyzed with a focus on the following features: (1) orientations of both binding sites for CXCL12, (2) orientations of the *N*-terminal domains of the two CXCL12 molecules and (3) stability of the complex (Table 3).

The analysis of the MD simulations for fr-HMGB1(I) revealed that both the domains are optimally oriented on the same side while the *N*-terminal domains are correctly oriented in the 61% of the analyzed conformations.

In fr-HMGB1(II) MD simulations, both domains and the *N*-terminal domains of CXCL12 were oriented in the same direction essentially for all the simulation time.

On the contrary, during the *MD* simulations of ds-HMGB1(II), which is the only conformation of the oxidized protein that can accommodate two CXCL12 molecules, both domains and the *N*-terminal of CXCL12 were oriented in opposite directions.

Furthermore, the protein tended to assume conformations in which BoxA and BoxB are close each other and then to be more compact (Figure S3).

In order to better assess the ability of the various heterocomplexes to trigger CXCR4 dimers, we determined the optimal distance between the CXCL12 N-terminal domains (44 Å, Figure S4) analyzing the X-ray structure of the CXCR4 dimer in complex with a viral chemokine (PDB ID code 4RWS[30], see methods). This value was then compared with the average distances measured in the MD simulations (Table 4).

For the fr-HMGB1(I) simulations the measured average value was approximatively 44.0 Å, while for the fr-HMGB1(II) simulations, the resulting value was larger than the reference value. However, a more accurate analysis of the simulations showed that the N-terminal domains stay at the proper distance during 2/3 of the simulation time.

In summary, MD simulations performed on the complexes obtained using molecular docking lead to some interesting observations. In fact, while fr-HMGB1 forms stable heterocomplexes with the N-terminal domains of CXCL12 optimally oriented for most of the time, all complexes between CXCL12 and ds-HMGB1, sampled in our simulations, are unstable and tend to assume conformations which are not competent for the binding to CXCR4 dimers.

Conclusions

Computational studies conducted on the two redox states of HMGB1 highlighted significant differences in the conformations adopted by the fr-HMGB1 and the ds-HMGB1 forms. In particular, RoG and SASA values computed for ds-HMGB1 were significantly lower than those of fr-HMGB1, indicating that the oxidized form of HMGB1 is more compact than the reduced one, while the local structure of BoxA remained essentially unchanged over 30 μs of MD simulations.

Cluster analysis and docking calculations provided insights into the molecular determinants underlying the enhancement of CXCR4 activation induced by the heterocomplex. In fact, the analysis of these structures showed that the ~75% of the conformations of fr-HMGB1 have BoxA and BoxB accessible for the binding of CXCL12. Furthermore, in these structures the two domains are optimally oriented to form $CXCL12_2$/HMGB1 heterocomplexes competent to bind and trigger CXCR4 dimers.

In conclusion, our computational studies support the hypothesis that the absence/presence of the disulfide bond in BoxA of HMGB1, regulates the formation of CXCL12/HMGB1 heterocomplex and the enhancement of CXCR4 signaling by the modulation of the HMGB1 conformational landscape.

These structural and dynamic information provide a better understanding of the mode of action of the $CXCL12_2$/HMGB1 heterocomplex, paving the way for novel approaches to design molecules able to interfere with its functions.

Supporting Information

Additional plots regarding RoG and SASA analysis; pictures of the stable compact conformations assumed by ds-HMGB1(II) with two CXCL12 molecules; representation of the CXCR4/vMIP-II complex; results of the residue-residue and atom-atom contact analysis; results of the cluster analysis carried out considering different cut-off levels. This material is available free of charge via the Internet at http://pubs.acs.org.

Acknowledgements

AC acknowledge the Swiss National Supercomputing Center (CSCS) for the availability of high-performance computing resources. This study was supported by grants from Krebsliga Schweiz (KLS-3839-02-2016-R) and the Swiss National Science Foundation (31003A-166472 to A.C and 3100A0-143718/1 to M.U.).

Bibliography

1. Kang, R.; Zhang, Q.; Zeh, H. J., 3rd; Lotze, M. T.; Tang, D., HMGB1 in cancer: good, bad, or both? *Clin. Cancer. Res.* 2013, 19, 4046-57.
2. Venereau, E.; De Leo, F.; Mezzapelle, R.; Careccia, G.; Musco, G.; Bianchi, M. E., HMGB1 as biomarker and drug target. *Pharmacol. Res.* 2016, 111, 534-44.
3. Naglova, H.; Bucova, M., HMGB1 and its physiological and pathological roles. *Bratisl. Lek. Listy.* 2012, 113, 163-71.
4. He, S.-J.; Cheng, J.; Feng, X.; Yu, Y.; Tian, L.; Huang, Q., The dual role and therapeutic potential of high-mobility group box 1 in cancer. *Oncotarget* 2017, 8, 64534-64550.
5. Andersson, U.; Yang, H.; Harris, H., High-mobility group box 1 protein (HMGB1) operates as an alarmin outside as well as inside cells. *Semin. Immunol.* 2018, 38, 40-48.
6. Andersson, U.; Tracey, K. J., HMGB1 is a therapeutic target for sterile inflammation and infection. *Annu. Rev. Immunol.* 2011, 29, 139-62.
7. Janko, C.; Filipović, M.; Munoz, L. E.; Schorn, C.; Schett, G.; Ivanović-Burmazović, I.; Herrmann, M., Redox Modulation of HMGB1-Related Signaling. *Antioxid. Redox Signal.* 2014, 20, 1075-1085.
8. Venereau, E.; Schiraldi, M.; Uguccioni, M.; Bianchi, M. E., HMGB1 and leukocyte migration during trauma and sterile inflammation. *Mol. Immunol.* 2013, 55, 76-82.
9. Bianchi, M. E.; Crippa, M. P.; Manfredi, A. A.; Mezzapelle, R.; Rovere Querini, P.; Venereau, E., High-mobility group box 1 protein orchestrates responses to tissue damage via inflammation, innate and adaptive immunity, and tissue repair. *Immunol. Rev.* 2017, 280, 74-82.
10. Andersson, U.; Yang, H.; Harris, H., Extracellular HMGB1 as a therapeutic target in inflammatory diseases. *Expert. Opin. Ther. Targets* 2018, 22, 263-277.
11. Yang, H.; Hreggvidsdottir, H. S.; Palmblad, K.; Wang, H.; Ochani, M.; Li, J.; Lu, B.; Chavan, S.; Rosas-Ballina, M.; Al-Abed, Y.; Akira, S.; Bierhaus, A.; Erlandsson-Harris, H.; Andersson, U.; Tracey, K. J., A critical cysteine is required for HMGB1 binding to Toll-like receptor 4 and activation of macrophage cytokine release. *Proc. Natl. Acad. Sci U. S. A.* 2010, 107, 11942-7.
12. Schrader, A. J.; Lechner, O.; Templin, M.; Dittmar, K. E.; Machtens, S.; Mengel, M.; Probst-Kepper, M.; Franzke, A.; Wollensak, T.; Gatzlaff, P.; Atzpodien, J.; Buer, J.; Lauber, J., CXCR4/CXCL12 expression and signalling in kidney cancer. *Br. J. Cancer* 2002, 86, 1250-6.
13. Schiraldi, M.; Raucci, A.; Munoz, L. M.; Livoti, E.; Celona, B.; Venereau, E.; Apuzzo, T.; De Marchis, F.; Pedotti, M.; Bachi, A.; Thelen, M.; Varani, L.; Mellado, M.; Proudfoot, A.; Bianchi, M. E.; Uguccioni, M., HMGB1 promotes recruitment of

inflammatory cells to damaged tissues by forming a complex with CXCL12 and signaling via CXCR4. *J. Exp. Med.* 2012, 209, 551-63.
14. Venereau, E.; Casalgrandi, M.; Schiraldi, M.; Antoine, D. J.; Cattaneo, A.; De Marchis, F.; Liu, J.; Antonelli, A.; Preti, A.; Raeli, L.; Shams, S. S.; Yang, H.; Varani, L.; Andersson, U.; Tracey, K. J.; Bachi, A.; Uguccioni, M.; Bianchi, M. E., Mutually exclusive redox forms of HMGB1 promote cell recruitment or proinflammatory cytokine release. *J. Exp. Med.* 2012, 209, 1519-28.
15. Lee, G.; Espirito Santo, A. I.; Zwingenberger, S.; Cai, L.; Vogl, T.; Feldmann, M.; Horwood, N. J.; Chan, J. K.; Nanchahal, J., Fully reduced HMGB1 accelerates the regeneration of multiple tissues by transitioning stem cells to G. *Proc. Natl. Acad. Sci U. S. A.* 2018, 115, 4463-4472.
16. Tirone, M.; Tran, N. L.; Ceriotti, C.; Gorzanelli, A.; Canepari, M.; Bottinelli, R.; Raucci, A.; Di Maggio, S.; Santiago, C.; Mellado, M.; Saclier, M.; François, S.; Careccia, G.; He, M.; De Marchis, F.; Conti, V.; Ben Larbi, S.; Cuvellier, S.; Casalgrandi, M.; Preti, A.; Chazaud, B.; Al-Abed, Y.; Messina, G.; Sitia, G.; Brunelli, S.; Bianchi, M. E.; Vénéreau, E., High mobility group box 1 orchestrates tissue regeneration via CXCR4. *J. Exp. Med.* 2018, 215, 303-318.
17. Cecchinato, V.; D'Agostino, G.; Raeli, L.; Nerviani, A.; Schiraldi, M.; Danelon, G.; Manzo, A.; Thelen, M.; Ciurea, A.; Bianchi, M. E.; Rubartelli, A.; Pitzalis, C.; Uguccioni, M., Redox-Mediated Mechanisms Fuel Monocyte Responses to CXCL12/HMGB1 in Active Rheumatoid Arthritis. *Front. Immunol.* 2018, 9. doi: 10.3389/fimmu.2018.0211
18. Abdulmahdi, W.; Patel, D.; Rabadi, M. M.; Azar, T.; Jules, E.; Lipphardt, M.; Hashemiyoon, R.; Ratliff, B. B., HMGB1 redox during sepsis. *Redox. Biol.* 2017, 13, 600-607.
19. Lv, W. L.; Arnesano, F.; Carloni, P.; Natile, G.; Rossetti, G., Effect of in vivo post-translational modifications of the HMGB1 protein upon binding to platinated DNA: a molecular simulation study. *Nucleic. Acids Res.* 2018, 46, 11687-11697.
20. Panneerselvam, S.; Durai, P.; Yesudhas, D.; Achek, A.; Kwon, H.-K.; Choi, S., Cysteine redox state plays a key role in the inter-domain movements of HMGB1: a molecular dynamics simulation study. *RSC Advances* 2016, 6, 100804-100819.
21. Antoine, D. J.; Harris, H. E.; Andersson, U.; Tracey, K. J.; Bianchi, M. E., A systematic nomenclature for the redox states of high mobility group box (HMGB) proteins. *Mol. Med.* 2014, 20, 135-7.
22. Fu, B.; Sahakyan, A. B.; Camilloni, C.; Tartaglia, G. G.; Paci, E.; Caflisch, A.; Vendruscolo, M.; Cavalli, A., ALMOST: an all atom molecular simulation toolkit for protein structure determination. *J. Comput. Chem.* 2014, 35, 1101-5.
23. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015, 11, 3696-3713
24. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, L. M., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 1983, 79, 926-935.
25. Joung, I. S.; Cheatham, T. E., Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B.* 2008, 112, 9020-9041.
26. Chiarelli, L. R.; Mori, M.; Barlocco, D.; Beretta, G.; Gelain, A.; Pini, E.; Porcino, M.; Mori, G.; Stelitano, G.; Costantino, L.; Lapillo, M.; Bonanni, D.; Poli, G.; Tuccinardi, T.; Villa, S.; Meneghetti, F., Discovery and development of novel salicylate synthase (MbtI) furanic inhibitors as antitubercular agents. *Eur. J. Med. Chem.* 2018,

155, 754-763.
27. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 1984, 81, 3684-3690.
28. Tom Darden, D. Y., and Lee Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys* 1993, 98, 10089-10092.
29. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical integration of the Cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J. Comp. Phys.* 1977, 23, 327-341.
30. Le Grand, S.; Götz, A. W.; Walker, R. C., SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comp. Phys. Commun.* 2013, 184, 374-380.
31. Roe, D. R.; Cheatham, T. E., PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 2013, 9, 3084-3095.
32. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J. Mol. Graph.* 1996, 14, 33-38.
33. Jörg, W.; S., S. P.; Clark, S. W., Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *J. Comp. Chem.* 1999, 20, 217-230.
34. Blau, C.; Grubmuller, H., g_contacts: Fast contact search in bio-molecular ensemble data. *Comp. Phys. Commun.* 2013, 184, 2856-2859.
35. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 2013, 29, 845-54.
36. Veldkamp, C. T.; Ziarek, J. J.; Su, J.; Basnet, H.; Lennertz, R.; Weiner, J. J.; Peterson, F. C.; Baker, J. E.; Volkman, B. F., Monomeric structure of the cardioprotective chemokine SDF-1/CXCL12. *Protein Sci.* 2009, 18, 1359-69.
37. van Zundert, G. C. P.; Rodrigues, J. P. G. L. M.; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A. M. J. J., The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* 2016, 428, 720-725.
38. Perez, J. J.; Tomas, M. S.; Rubio-Martinez, J., Assessment of the Sampling Performance of Multiple-Copy Dynamics versus a Unique Trajectory. *J. Chem. Inf. Mod.* 2016, 56, 1950-1962.
39. Qin, L.; Kufareva, I.; Holden, L. G.; Wang, C.; Zheng, Y.; Zhao, C.; Fenalti, G.; Wu, H.; Han, G. W.; Cherezov, V.; Abagyan, R.; Stevens, R. C.; Handel, T. M., Structural biology. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science* 2015, 347, 1117-22.
40. Wang, J.; Tochio, N.; Takeuchi, A.; Uewaki, J.; Kobayashi, N.; Tate, S., Redox-sensitive structural change in the A-domain of HMGB1 and its implication for the binding to cisplatin modified DNA. *Biochem. Biophys. Res. Commun.* 2013, 441, 701-6.
41. Xavier, D.; Karl, G.; Bernhard, J.; Dieter, S.; F., v. G. W.; E., M. A., Peptide Folding: When Simulation Meets Experiment. *Angew. Chem. Int. Ed.* 1999, 38, 236-240.

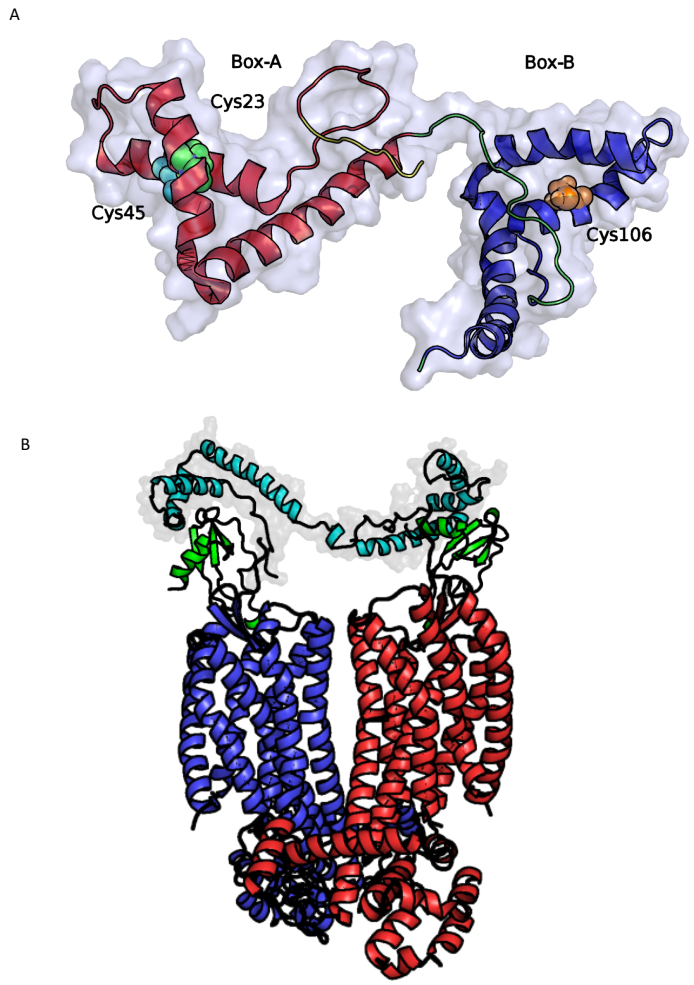


Figure 1. (A) Structure of HMGB1 (PDB ID code 2YRQ) solved by NMR. Protein domains are presented in different colors: BoxA (red), BoxB (blue), and the loop between the two domains (green). The three cysteines located at positions 23 and 45 in BoxA and 106 in BoxB are displayed as van der Waals balls in different colors. (B) Explicative representation, of the binding of the $CXCL12_2$/HMGB1 to a CXCR4 dimer. HMGB1 is depicted in aquamarine, the two CXCL12 in green, while the two CXCR4 monomers in blue and red respectively.

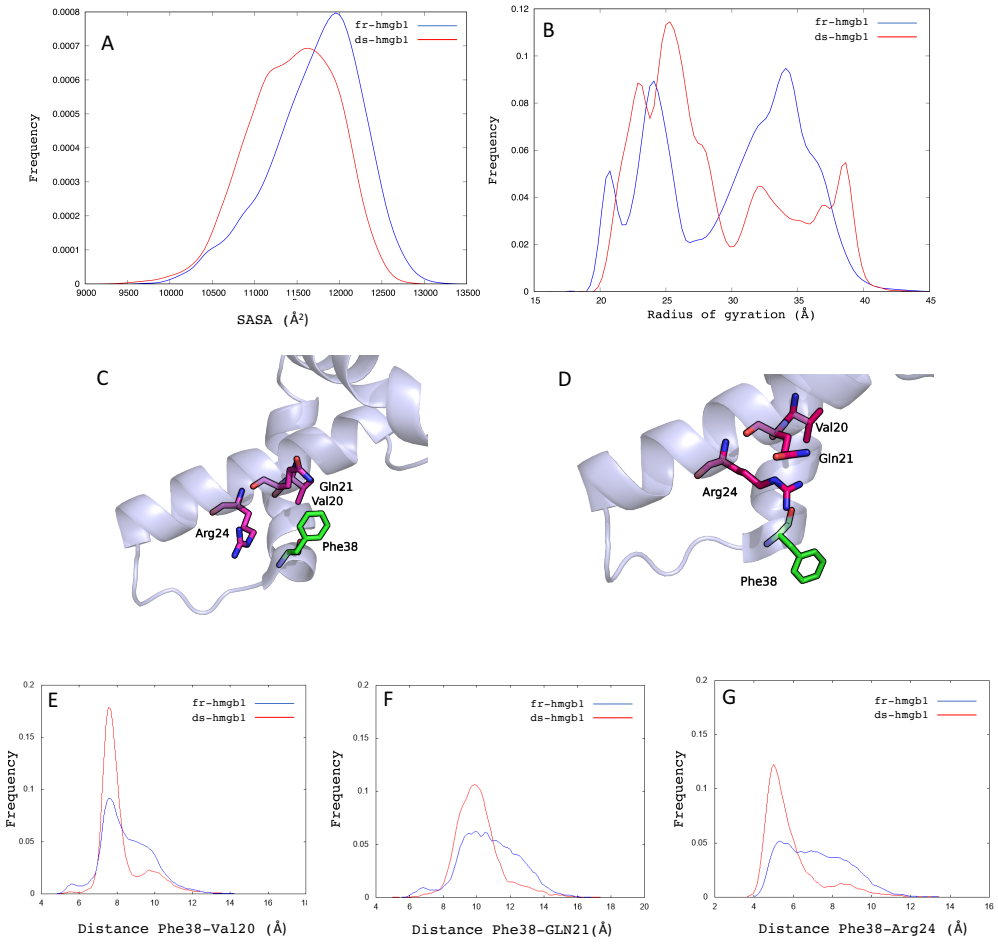


Figure 2. (A) Histograms of the solvent accessible surface area (SASA) and (B) radius of gyration (RoG) computed using all residues of the protein. Details about the Phe38 orientation in the ds- (C, pdb code 2RTU) and fr- (D, 2YRQ) HMGB1. Histograms of the distance between the center of mass (COM) of Phe38 and COM of Val20 (E), Gln21 (F) and Arg24 (G). In all histograms, the data for fr-HMGB1 are shown in blue while those of ds-HMGB1 in red.

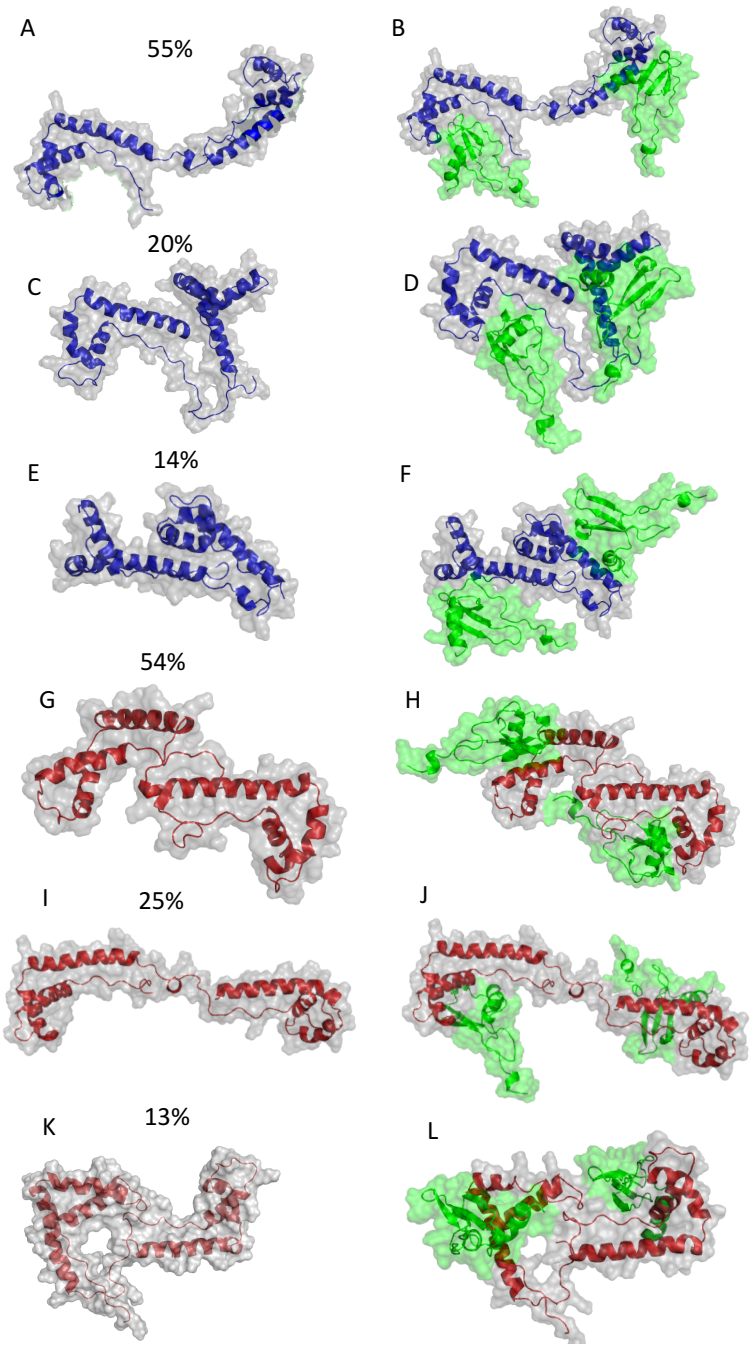


Figure 3. Representative conformations of the three most populated clusters (cluster centers) of fr-HMGB1 (A, C, and E) and ds-HMGB1 (G, I, and K) obtained from the cluster analysis performed using the GROMOS method.[41] The cluster size is reported as a percentage of the entire conformational ensemble. Structures of the complexes between the three most representative fr-HMGB1 (B, D, and F) and ds-HMGB1 (H, J, and L) conformations and two CXCL12 molecules (green) were obtained using protein-protein docking software HADDOCK.

Table 1. Summary of the MD simulations2qz performed in this study.

System	Description	Simulation Time
fr-HMGB1	HMGB1 NMR structure (PDB ID code 2YRQ)	30x1 μs
ds-HMGB1	HMGB1 with a disulfide bond between Cys23-Cys45	30x1 μs
fr-HMGB1(I)	First representative cluster of the fr-HMGB1 + $CXCL12_2$ complex	3x500 ns
fr-HMGB1(II)	The second representative cluster of the fr-HMGB1 + $CXCL12_2$ complex	3x500 ns
ds-HMGB1(II)	The second representative cluster of the ds-HMGB1 + $CXCL12_2$complex	3x500 ns

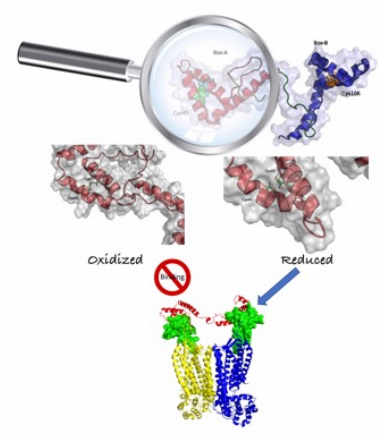
Table 2. Residues involved in the interaction between HMGB1 and CXCL12 used to drive the docking procedure.

Residues of HMGB1 interacting with CXCL12	
BoxA	14S, 16Y, 17A, 20V, 24R, 25E, 38F, 39S, 50K, 57K
BoxB	96K, 103F, 104L, 113I, 114K, 115G, 116E, 120L, 137A, 155Y, 158D
Residues of CXCL12 interacting with HMGB1	
With BoxA	18V, 19A, 23V, 24K, 25H, 38I, 40A, 41R, 42L, 44N, 48Q, 59N, 64K
With BoxB	15E, 18V, 20R, 23V, 24K, 38I, 40A, 41R, 45N, 51I, 57W, 58I, 59N, 64K, 65A, 66L

Table 3. Percentage of the frames sampled in the MD simulations in which the two HMGB1 domains BoxA and BoxB or the two N-terminal domains of CXCL12 have the same orientation.

fr-HMGB1(I)		fr-HMGB1(II)		ds-HMGB1(II)	
Domains	94%	Domains	100%	Domains	0%
NT-ends	61%	NT-ends	92%	NT-ends	0%

Table 4. Distance between K1 of the two CXCL12 molecules in complex with HMGB1 measured during MD simulations. The distance was only measured in simulations in which the two N-terminal domains are properly oriented to trigger CXCR4 dimers.

Distance between K1 of CXCL12$_2$ molecules (*Ref. = 44.0 Å)*			
	Sim 1	Sim 2	Sim 3
fr-HMGB1(I)	-	44.66 ± 14.57 Å	48.66 ± 15.03 Å
fr-HMGB1(II)	53.25 ± 15.40 Å	56.42 ± 10.81 Å	48.88 ± 13.03 Å
ds-HMGB1(II)	-	-	-

Graphical table of content